7-2020

RIFT'ing the Waves: Developing and applying an algorithm to infer properties of gravitational wave sources

Jacob Lange
jal7686@rit.edu

RIFT'ing the Waves: Developing and applying an algorithm to infer properties of gravitational wave sources

---

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the degree of Ph.D. *Doctor of Philosophy*

in Astrophysical Sciences and Technology

Jacob Lange

School of Physics and Astronomy

Rochester Institute of Technology

Rochester, New York

July 2020

ASTROPHYSICAL SCIENCES AND TECHNOLOGY
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of *Jacob Lange* has been examined and approved by the dissertation committee as satisfactory for the dissertation requirement for the Ph.D. degree in Astrophysical Sciences and Technology.

Dr. Nathan Cahill, Committee Chair

Dr. Carlos Lousto, Committee Member

Dr. Geoffrey Lovelace, Committee Member

Dr. Richard O'Shaughnessy, Dissertation Advisor

Date

RIFT'ing the Waves: Developing and applying an algorithm to infer properties of gravitational wave sources

By

Jacob Lange

A dissertation submitted in partial fulfillment of the requirements for the degree of Ph.D. in Astrophysical Sciences and Technology, in the College of Science, Rochester Institute of Technology.

July, 2020

Approved by

Dr. Andrew Robinson	Date
Director, Astrophysical Sciences and Technology	

Declaration

I, JACOB A. LANGE ("the Author"), declare that no part of this dissertation is substantially the same as any that has been submitted for a degree or diploma at the Rochester Institute of Technology or any other University. I further declare that Chapter 2 is entirely my own; Chapters 3 and sections therein are modified reproductions of work in which I am a co-primary author, and include contributions from other authors, who are credited below. Chapter 4 includes work I produced for the LIGO Scientific Collaboration and the Virgo Collaboration of which I am apart of. Those who have contributed scientific or other collaborative insights are fully credited in this dissertation, and all prior work upon which this dissertation builds is cited appropriately throughout the text. This dissertation was successfully defended in Rochester, NY, USA on the 24^{th} of July 2020.

Modified portions of this dissertation will be published by the Author in a peer-reviewed paper appearing in Phys. Rev. D(PRD):

- **Chapter 2** is based on the paper [1], entitled *Rapid and accurate parameter inference for coalescing compact binaries.* (2019, submitted to PRD) co-authored by Richard O'Shaughnessy (the dissertation advisor) and Monica Rizzo.

Modified portions of this dissertation have also been published by the Author and the other co-authors in peer-reviewed papers appearing in Phys. Rev. D(PRD) and Astrophysical Journal(ApJ):

- **Section 3.2** is based on the paper [2], entitled *Systematic challenges for future gravitational wave measurements of precessing binary black holes.* (2017, PRD, 96, 124041)

co-authored by Andrew R. Williamson (first author), Richard O'Shaughnessy (the dissertation advisor), James A. Clark, Prayush Kumar, Juan Calderón Bustillo, and John Veitch

- **Section 3.3** is based on the paper [3], entitled *Targeted numerical simulations of binary black holes for GW170104* (2018, PRD, 97, 064027) co-authored by James Healy (first author), Richard O'Shaughnessy (the dissertation advisor), Carlos O. Lousto (dissertation committee member), Manuela Campanelli, Andrew R. Williamson, Yosef Zlochower, Juan Calderón Bustillo, James A. Clark, Chris Evans, Deborah Ferguson, Sudarshan Ghouge, Karan Jani, Bhavesh Khamesra, Pablo Laguna, Deirdre M. Shoemaker, Michael Boyle, Alyssa García, Daniel A. Hemberger, Lawrence E. Kidder, Prayush Kumar, Geoffrey Lovelace (dissertation committee member), Harald P. Pfeiffer, Mark A. Scheel, and Saul A. Teukolsky

- **Section 3.4** is based on the paper [4], entitled *Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems* (2020, PRD, 101, 124054) co-authored by Feroz H. Shaik (first author), Scott E. Field, Richard O'Shaughnessy (the dissertation advisor), Vijay Varma, Lawrence E. Kidder, Harald P. Pfeiffer, and Daniel Wysocki

- **Section 3.5** is based on the paper [5], entitled *Inferring parameters of GW170502 - the loudest intermediate mass black hole trigger in LIGO's O1/O2 data* (2020, accepted to ApJ) co-authored by Richard Udall (first author), Karan Jani, Richard O'Shaughnessy (the dissertation advisor), James A. Clark, Laura Cadonati, Deirdre Shoemaker, and Kelly Holley-Bockelmann

Results included in this text will also be published by the LIGO Scientific Collaboration and the Virgo Collaboration, which the Author is apart of, in peer-reviewed papers appearing in Phys. Rev. D(PRD) and/or Phys. Rev. X(PRX):

- **Chapter 4** includes the same results as papers [6, 7], entitled *Properties and astrophysical implications of the 150 $M_\odot$ binary black hole merger GW190521.1* (2020, submitted

to PRD) and *GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run* (2020, planning to submit to PRX) both authored by the LIGO Scientific Collaboration and the Virgo Collaboration.

To my parents: Being your son is my greatest honor.

Acknowledgments

We are grateful for computational resources provided by the Leonard E Parker Center for Gravitation, Cosmology and Astrophysics at University of Wisconsin-Milwaukee; by the Albert Einstein Institute in Hannover, Germany; by California Institute of Technology; by the LIGO Livingston Computing Cluster; and by the LIGO Hanford Computing Cluster via the LIGO Data Grid (LDG) cluster. We are grateful for computational resources provided by Cardiff University, and funded by an STFC grant supporting UK Involvement in the Operation of Advanced LIGO. We also acknowledge the use of IUCAA LDG cluster Sarathi for the computational/numerical work. We are grateful for computational resources provided by the CARNiE cluster at UMassD, which is supported by the ONR/DURIP Grant No. N00014181255. We are also grateful for the computational resources provided by the Open Science Grid. We would like to thank the Rochester Institute of Technology, Simulation eXtreme Spacetimes Collaboration, and Georgia Tech numerical relativity groups for supplying the waveforms used in the work. We would like to thank Rochester Institute of Technology for allowing us to complete this work.

On a personal note, I would also like to thank my PhD committee for the time they took with my defense. I would also like to thank my advisor Richard O'Shaughnessy for guiding and mentoring me throughout the competition of this work. I would also like to thank all the professors at Rochester Institute of Technology who taught my classes and gave me guidance. I especially would like to thank Andy Robinson for guiding me through the entire PhD program. I especially would also like to thank Carlos Lousto and Jim Healy for collaborating with me and guiding me. I am grateful for all the long hours I spent with my fellow graduate students

doing homework, discussing topics, or just having fun; I especially would like to thank Triana Almeyda who has always supported and encouraged me. Most of all, I would like to thank my parents who has always believed in and supported me.

Abstract

With the Advanced LIGO and Virgo ground-based detectors consistently identifying more compact binary coalesces, the need for fast, reliable, and unbiased parameter inference is ever more vital. To that end, we introduce RIFT: an algorithm to perform Rapid parameter inference on gravitational wave sources via Iterative FiTting. To demonstrate RIFT can recover the correct parameters of coalescing compact binary systems, we compare results to the well-tested LALInference parameter inference software. We provide several examples where the unique speed and flexibility of RIFT enables otherwise intractable or awkward parameter inference analyses, such as (a) adopting costly and novel models for outgoing gravitational waves and (b) mixed-model result, each suitable to different parts of the compact binary parameter space and allowing one to use more sophisticated approximations where valid but still producing a complete posterior distribution. We also demonstrate how RIFT can be applied specifically to binary neutron stars, both for parameter inference and direct constraints on the nuclear equation of state.

We also show that two precessing models often used in inferring the properties of coalescing black hole binaries disagree substantially when sources have modestly large spins and modest mass ratios. We demonstrate these disagreements using standard figures of merit and the parameters inferred for some detections of binary black holes from O1 and O2. By comparing to numerical relativity, we confirm these disagreements reflect systematic errors. We provide concrete examples to demonstrate that these systematic errors can significantly impact inferences about astrophysically significant binary parameters.

In response to LIGO's observation of GW170104, a series of full numerical simulations of binary black holes were performed, each designed to replicate likely realizations of its dynamics and radiation. These simulations have been performed at multiple resolutions and with two independent techniques to solve Einstein's equations. For both the nonprecessing and precessing simulations, we demonstrate the two techniques agree at a precision substantially in excess of statistical uncertainties in current LIGO's observations. Conversely, we demonstrate

that these full numerical solutions contain information which is not accurately captured with the approximate phenomenological models. To quantify the impact of these differences on parameter inference for GW170104 specifically, we compare the predictions of our simulations and these approximate models to LIGO's observations of GW170104.

Using one of the novel numerical relativity surrogate models, we also investigate the importance of higher order modes when inferring the parameters of coalescing compact binaries. We focus on examples relevant to the current three-detector network of observatories with a detector-frame mass set to $120M_{\odot}$ and with signal amplitudes values that are consistent with plausible candidates for the next few observing runs. We show that for such systems the higher mode content will be important for interpreting coalescing binary black holes, reducing systematic bias, and computing properties of the remnant object.

Using similar tools, we finally use RIFT to analyze many real data events. This includes the loudest marginal intermediate mass binary black hole trigger from the 1st and 2nd Observing Runs as well as a subset of the events from the first half of the 3rd Observing Run. This includes both 15 binary black hole candidates and 1 binary neutron star candidate.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since its first detection on September 14, 2015 of gravitational waves (GW) from a binary black hole (BBH) system [19], the LIGO-Virgo Collaboration (LVC) continued to search and infer the properties of new GW sources. Over the upcoming years, the LIGO (Laser Interferometer Gravitational Wave Observatory) [20] and Virgo observatories [21] will go through phases of observation and maintenance. During these maintenance periods, the observatories will have their sensitivity increase that will in turn increase the number of detections. Over the coming years, hundreds more events will be confidently identified with proportionally more less-confident candidates at the margins of the observatories sensitivities. These plethora of events provide an opportunity and a challenge. On one hand, it will be essential to obtain reliable and unbiased parameter estimation (PE). These will be vital to achieve the full potential of high-precision tests of general relativity; constraints on nuclear matter; and inferences about present-day populations and progenitor astrophysics. On the other hand, these results will be stable only with a careful exploration of model systematics, which at present can be difficult with conventional inference tools due to computational cost especially with low-mass systems.

My work presented in this dissertation builds off of previous work [22] by developing a highly-parallelizable grid-based PE strategy. Using this strategy, each GW candidate is compared to a grid of points in source parameter space producing a range of likelihood values. After calculating a likelihood and marginalizing over the extrinsic parameters ($\ln \mathcal{L}_{\rm marg}$), the

continuous $\ln \mathcal{L}_{\rm marg}$ distribution can be constructed by interpolation. Using this continuous $\ln \mathcal{L}_{\rm marg}$ distribution and assuming some prior, the posterior distribution for intrinsic parameters can be obtained, see [18] and Chapter 2 for more details.

In this chapter, I briefly review basic concepts of gravitational radiation and parameter inference as well as describe the different waveform families and the corresponding models used in this dissertation. I then briefly describe the previous studies that this work is derived from and layout the organization of the rest of the dissertation.

1.1 Bayesian Inference of Gravitational Wave Sources

1.1.1 Bayesian Inference

Most codes use some variation of Bayesian Inference to estimate the properties of gravitational wave sources, which can describe the probability density of some unknown parameter or describe the current knowledge of some given hypothesis [23, 18, 1, 22]. Given two hypotheses about an observation A and B, the joint probability, the probability of both A and B being true, is $P(A, B)$. This joint probability can be expressed in terms of the conditional probability and the individual probability of the hypotheses. If $P(B)$ is the individual probability that B is true and $P(A|B)$ is the conditional probability that A is true given B is true, the joint probability can be defined as follows:

$$P(A, B) = P(B)P(A|B). \tag{1.1.1}$$

Similarly if $P(A)$ is the individual probability that A is true and $P(B|A)$ is the conditional probability that B is true given BA is true, the joint probability can be defined as follows:

$$P(A, B) = P(A)P(B|A). \tag{1.1.2}$$

Setting Eq. 1.1.1 and 1.1.2 equal to each other and solving for one of the conditional probabilities, we arrive a Bayes' Theorem:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \tag{1.1.3}$$

This is true for not only the probability but also the probability density where $\int p(A)dA = 1$:

$$p(A|B) = \frac{p(A)p(B|A)}{p(B)}. \tag{1.1.4}$$

In this equation, $p(A|B)$ is considered the *posterior probability density*, $p(B|A)$ is often called the *likelihood*, $p(A)$ is called the *prior probability* of A, and $p(B)$ is often called the *evidence* and is a normalization constant with the numerator as the integrand (Note that since A and B are arbitrary hypotheses, one could also solve for $p(B|A)$). If we assume A are the unknown parameters $\vec{\gamma} = \{\gamma_1 ... \gamma_N\}$ and B is the gravitational wave data, we can write Eq. 1.1.3 as:

$$p(\vec{\gamma}|d) = \frac{p(\vec{\gamma})p(d|\vec{\gamma})}{p(d)} \tag{1.1.5}$$

If we define the H as the state of knowledge about an uncertain hypothesis with probability $P(H) \in [0, 1]$ and $\int p(\gamma|H)d\gamma = 1$, we can rewrite Eq. 1.1.5 to:

$$p(\vec{\gamma}|d, H) = \frac{p(\vec{\gamma}|H)p(d|\vec{\gamma}, H)}{p(d|H)} \tag{1.1.6}$$

where $p(\vec{\gamma}|d, H)$ describes the joint probability density of a multidimensional parameter space that describes the knowledge of all parameters and their relationship.

A set of data $d_k(t)$ collected from a GW detector is typically decomposed into the inherent noise of the detect and the signal of the GW:

$$d_k(t) = h_k(t) + n_k(t). \tag{1.1.7}$$

We define a weighted inner-product of two complex Fourier-domain functions $\tilde{a}(f), \tilde{b}(f)$ with some weighting function $S(f)$:

$$\langle a|b\rangle_k \equiv \int_{-\infty}^{\infty} 2df \frac{\tilde{a}(f)^* \tilde{b}(f)}{S_{h,k}(|f|)}. \tag{1.1.8}$$

In practice we adopt a low-frequency cutoff $f_{\rm min}$ so all inner products are modified to

$$\langle a|b\rangle_k \equiv 2 \int_{|f|>f_{\rm min}} df \frac{[\tilde{a}(f)]^* \tilde{b}(f)}{S_{h,k}(|f|)}. \tag{1.1.9}$$

If the detector noise is assuming to be Gaussian, one can write the probability of some set of noise realizations in each detector as:

$$p(d|H_0) \propto \prod_k \exp\Big(-\frac{\langle n|n\rangle_k}{2}\Big) \tag{1.1.10}$$

where H_0 indicates the hypothesis that the data is only Gaussian noise. With a GW signal present, the data is the noise plus some response to the signal $h(\vec{\gamma})$. The probability of some measured d given the presence of the signal with parameter values $\vec{\gamma}$ is

$$p(d|\vec{\gamma}, H_1) \propto \prod_k \exp\Big(-\frac{\langle d-h(\vec{\gamma})|d-h(\vec{\gamma})\rangle_k}{2}\Big) \tag{1.1.11}$$

where H_1 is the hypothesis that the data consists of Gaussian noise plus a signal. Using Eq. 1.1.6, the posterior probability of the parameter $\vec{\gamma}$ under hypothesis H_1 is

$$p(\vec{\gamma}|d, H_1) = \frac{p(\vec{\gamma}|H_1)p(d|\vec{\gamma}, H_1)}{p(d|H_1)} \tag{1.1.12}$$

where

$$p(d|H_1) = \int p(d|\vec{\gamma}, H_1)p(\vec{\gamma}|H_1)d\vec{\gamma} \tag{1.1.13}$$

We now introduce the likelihood ratio:

$$\mathcal{L}(d|\vec{\gamma}) = \prod_k \frac{\exp\{-\langle d-h(\vec{\gamma})|d-h(\vec{\gamma})\rangle_k/2\}}{\exp\{-\langle d|d\rangle_k/2\}} \tag{1.1.14}$$

and can rewrite equation 1.1.12 as

$$p(\vec{\gamma}|d, H_1) = \frac{\mathcal{L}(\vec{\gamma}|d)p(\vec{\gamma})}{Z} \tag{1.1.15}$$

where Z can be be called the Bayes Factor defined as:

$$Z = \int \mathcal{L}(\vec{\gamma}|d)p(\vec{\gamma})d\vec{\gamma} \equiv \frac{p(d|H_1)}{p(d|H_0)}. \tag{1.1.16}$$

See [22] for more details. This shows how much more likely the observed data is under one model over the other.

1.1.2 Gravitational Wave Parameters

For the purposes of this dissertation, we will emphasize the segregation of the GW parameters $\vec{\gamma}$ into its intrinsic parameters $\vec{\lambda}$, which describe the fundamental properties of the binary, and its extrinsic parameters $\vec{\theta}$, which describe it's orientation and location in space and the time it arrives at the detector:

$$\vec{\gamma} = \vec{\lambda} + \vec{\theta} \tag{1.1.17}$$

The following list gives a brief description of the intrinsic parameters $\vec{\lambda}$ used to characterize a GW signal:

- For our purposes, we will re-parameterize the individual masses m_1 and m_2 (where the common convention $m_1 \geq m_2$) into the chirp mass defined as:

$$\mathcal{M} = (m_1 m_2)^{3/5}(m_1 + m_2)^{-1/5} \tag{1.1.18}$$

  and the symmetric mass ratio defined as:

$$\eta = \frac{(m_1 m_2)}{(m_1 + m_2)^2}. \tag{1.1.19}$$

Other mass parameters that are also used in this dissertation are the total mass

$$M = m_1 + m_2 \tag{1.1.20}$$

and mass ratio defined as either

$$q = m_1/m_2 \text{ or } q = m_2/m_1 \tag{1.1.21}$$

Note that in Chapters 2,4 and Sections 3.3, 3.5 defines $q \leq 1$ or $q = m_2/m_1$ but Sections 3.2,3.4 defines $q \geq 1$ or $q = m_1/m_2$.

- If the objects in the binary have any spin, it is useful to define the effective spin [24, 25, 26], which is a combination of the spin components along the orbital angular momentum, in the following way,

$$\chi_{\rm eff} = (\vec{S_1}/m_1 + \vec{S_2}/m_2) \cdot \hat{\mathbf{L}}/M \tag{1.1.22}$$

where $\vec{S_1}$ and $\vec{S_2}$ are the spin vectors of each object, related to the dimensionless spins $\vec{\chi_i}$ on the individual compact object via

$$\vec{\chi_i} = \vec{S_i}/m_i^2. \tag{1.1.23}$$

We will express the dimensionless spins in terms of cartesian components $\chi_{i,x}, \chi_{i,y}, \chi_{i,z}$, expressed relative to a frame with $\hat{\mathbf{z}} = \hat{\mathbf{L}}$

- If the binary's objects have any tides, we characterize the tidal deformability of each compact object by a dimensionless parameter Λ_i, which is zero for black holes (BH). The leading-order effects of tidal deformation enter into the gravitational wave signal

via two quantities:

$$\widetilde{\Lambda} = \frac{16}{13}\frac{(m_1+12m_2)m_1^4\Lambda_1 + (m_2+12m_1)m_2^4\Lambda_2}{(m_1+m_2)^5}$$
$$= \frac{8}{13}\left[(1+7\eta-31\eta^2)(\Lambda_1+\Lambda_2) + \sqrt{1-4\eta}(1+9\eta-11\eta^2)(\Lambda_1-\Lambda_2)\right] \quad (1.1.24)$$

$$\delta\widetilde{\Lambda} = \frac{1}{2}\left[\sqrt{1-4\eta}\left(1-\frac{13272}{1319}\eta+\frac{8944}{1319}\eta^2\right)(\Lambda_1+\Lambda_2) + \left(1-\frac{15910}{1319}\eta+\frac{32850}{1319}\eta^2+\frac{3380}{1319}\eta^3\right)(\Lambda_1-\Lambda_2)\right], \quad (1.1.25)$$

In most cases of current astrophysical interest, $\widetilde{\Lambda}$ can be weakly constrained and $\delta\widetilde{\Lambda}$ cannot be constrained at all.

The following list gives a brief description of the extrinsic parameters $\vec{\theta}$ used to characterize a GW signal:

- $t_{\rm geo}$ is the time at which the coalescing point of the binary's waveform arrives at the geocenter of Earth.
- α and δ are the right ascension and declination.
- θ_{JN} is the inclination between the system's total angular momentum with respect to the line of sight of Earth.
  - ι is the inclination between the system's orbital angular momentum and the line of sight of Earth. For non-precessing systems, this parameter is identical to θ_{JN}.
- d_L is the luminosity distance.
- ψ is the polarization angle.
- ϕ_c is the orbital phase of the binary at coalescence.

1.2 Waveform Models

A number of waveform model families have been developed to emulate the outgoing gravitational radiation from compact binary coalescences. These models can vary in the physics they

simulate, the approximations assumed, the computational costs, the domain in which the waveforms are generated (time- or frequency-domain), and their parameter validity regions. These models take input parameters $\vec{\gamma}$ and outputs the components of the strain in the kth detector $h_{+,k}$ and $h_{\times,k}$ (or for our purposes h_{lm}s). In the work presented here, we focus on 4 major waveform families: the inspiral-merger-ringdown phenomenological models (IMRPhenom), the effecive-one-body models (EOB), the numerical relativity simulations (NR), and the NR surrogate models (NRSur).

The EOB approach models the inspiral and spin dynamics of coalescing binaries via an ansatz for the two-body Hamiltonian [27], whose corresponding equations of motion are numerically solved in the time domain. For nonprecessing binaries, outgoing gravitational radiation during the inspiral phase is generated using an ansatz for resumming the post-Newtonian expressions for outgoing radiation including non-quasicircular corrections, for the leading-order $\ell = 2$ subspace. For the merger phase of nonprecessing binaries, the gravitational radiation is generated via a resummation of many quasinormal modes, with coefficients chosen to ensure smoothness. The final BH's mass and spin, as well as some parameters in the nonprecessing inspiral model, are generated via calibration to numerical relativity simulations of BBH mergers. For precessing binaries, building off the post-Newtonian ansatz of separation of timescales and orbit averaging [28, 29, 30, 31], gravitational radiation during the inspiral is modeled as if from an instantaneously nonprecessing binary (with suitable nonprecessing spins), in a frame in which the binary is not precessing [32, 33, 34]. During the merger, the radiation is approximated using the same final BH state, with the same precession frequency.[1] With well-specified initial data in the time domain, this method can be directly compared to the trajectories [37] and radiation [11] of numerical BBH spacetimes. In this work we use SEOBNRv4, a model for nonprecessing binaries [38]; SEOBNRv4_ROM, a fast surrogate model for SEOBNRv4 [38]; and SEOBNRv4P (and the older version SEOBNRv3), a model for precessing binaries [39, 11]. The non-precessing models only contain the dominant (2,2) mode, and the precessing models include the $(2, \pm 2); (2, \pm 1)$ modes in the co-precessing frame.

[1] This choice of merger phase behavior is known to be inconsistent with precessional dynamics during merger [35, 36].

For any higher order mode (HOM) analysis, we use either the nonspinning EOBNRv2HM [40, 41] or the nonprecessing SEOBNRv4HM [42], which both include the modes: $(2, \pm 2); (2, \pm 1); (3, \pm 3); (4, \pm 4); (5, \pm 5)$; or the fully precessing version SEOBNRv4PHM [43, 11] (which includes the same number of modes in the co-precessing frame).

The IMRPhenomPv2 model is a part of an approach that attempts to approximate the leading-order gravitational wave radiation using phenomenological fits to the Fourier transform of this radiation, computed from numerical relativity simulations and post-newtonian calculation [44, 45, 46]. Also using information about the final BH state, this phenomenological frequency-domain approach matches standard approximations for the post-Newtonian gravitational wave phase to an approximate, theoretically-motivated spectrum characterizing merger and ringdown. For IMRPhenomPv2, precession is also incorporated by a "co-rotating frame" ansatz, here implemented via a stationary-phase approximation to the time-domain rotation operations performed for SEOBNRv3 and SEOBNRv4P. The IMRPhenomD model [12] is the nonprecessing limit of IMRPhenomPv2 (both only contain the dominant (2,2) mode). We leave the use of the more recently developed sophisticated models for future studies: IMRPhenomHM [47], IMRPhenomPv3 [48], IMRPhenomPv3HM [49], IMRPhenomXAS [50],IMRPhenomXP [51, 52], and IMRPhenomXPHM [51, 52].

Surrogate models provide efficient and accurate representations of the gravitational wave strain, by interpolating between evaluations of a costly reference model. They have been applied to long duration signals [53, 54], arbitrarily many harmonic modes [53, 55], spinning binary systems [54, 56], and neutron star models with tidal effects [56]. We focus on the NRSur that directly interpolates a large set of both precessing and non-precessing NR simulations. In this work, we use surrogates developed to reproduce multimodal radiation from non-precessing [57] as well as precessing binary systems [58, 36, 59]. These surrogates are demonstrably much more accurate in their domain of validity than the approximations described above. While the non-precessing NRSur model NRHybSur3dq8, used in Section 3.4 and 3.5, is hybridized to allow for arbitrary long waveforms, the precessing surrogates NRSur7dq2 [36], used in Chapter 2, and the update version NRSur7dq4, used in Section 3.5 and

Chapter 4, have a finite duration of ~ 20 cycles due to the same finite duration of the NR simulations involved in the training set. Also due to the limited duration and parameter space coverage, each of these surrogates have a limited range of validity. As the names suggest, the mass ratio restrictions are $q < 2$ for NRSur7dq2, $q < 4$ for NRSur7dq4, and $q < 8$ for NRHybSur3dq8. The spins of the individual black holes are restricted to $|\chi_i| < 0.8$ for all surrogates. These spins restrictions are eased to varying degrees in the different chapters to follow. The precessing surrogates include all $\ell \leq 4$ modes, and the non-precessing surrogate includes all $\ell \leq 4$ and $(5, \pm 5)$ modes except the $(4, \pm 1)$ and $(4, 0)$ modes.

Models for binary neutron stars (BNS) account for the response of each neutron star's structure into their estimate for dynamics of and radiation during the inspiral and coalescence. Using frequency domain methods which generalize classic post-Newtonian calculations [60, 61], Dietrich et al [62] introduced a simple way to add leading-order tidal effects to the nonprecessing models described above. Recently Dietrich et al [63] implemented this effect into SEOBNRv4_ROM and IMRPhenomD, which we denote by the postfix _NRTidal. These _NRTidal approximations were developed and calibrated using an effective one body model (TEOBResumS) that incorporates the effects of adiabatic tides and spin [64, 65]; see [63] and references therein. Another model (SEOBNRv4T) and it's frequency-domain surrogate model (SEOBNRv4T_surrogate) approximates binary inspiral and neutron star dynamics by allowing neutron stars to have both spin and dynamical tides [38, 66, 67, 68, 69]. As described in part in [63], these approximations include tidal effects at differing levels of completeness. For example, among the models described here only TEOBResumS and SEOBNRv4T incorporate the quadrupole-monopole interaction [70] into the dynamics; only TEOBResumS includes higher modes into the outgoing radiation; and only SEOBNRv4T uses dynamic tides. While none of these models can account for neutron star spins that are misaligned from the orbital angular momentum, these models capture the leading-order features most relevant for small neutron star (NS) spins.

In the 1970s and 1980s as the convention tools for studying relativity were reaching their limit, scientist researched more modern ways to increase their understanding. One of these

ways was solving for a numerical solution to Einstein's equations using supercomputers. While there were some initial challenges to the numerical evolution codes, the breakthroughs [71, 72, 73] in numerical relativity allowed for detailed predictions for the gravitational waves from the late inspiral, plunge, merger and ringdown of black hole binary systems. Today, there are a number of different NR groups that produce catalogs of different NR waveforms [74, 75, 76, 77, 78, 79]. These can be used not only for determining how the individual masses and spins of the orbiting binary relate to the properties of the final remnant produced after merger but also for direct comparison to the GW data for PE (see [18] and Section 1.3 for a summary of that work).

The Rochester Institute of Technology (RIT) group evolved the BBH data sets described below using the LazEv [80] implementation of the moving puncture approach [72, 73]. The LazEv code uses the EinsteinToolkit [81, 82] / Cactus [83] / Carpet [84] infrastructure. Similarly, the Georgia Tech (GT) [75] simulations were obtained with the Maya code [85, 86, 87, 88, 89, 90, 91, 92], which is also based on the BSSN formulation with moving punctures. Finally, the Simulating eXtreme Spacetimes Collaboration (SXS) simulations are carried out using the Spectral Einstein Code (SpEC) [93, 74, 78]. These codes were used to generated NR follow up simulations in which the work Section 3.3 was apart of. While not emphasized here, the larger NR catalogs are used in tandem with RIFT to analyze GW signals by comparing directly to the data. See Section 4.2.6 and publications [18, 6, 7] for more details and examples.

1.3 Related Work

1.3.1 Efficiently evaluating the Likelihood

This section summarizes the methodology of segregating the intrinsic and extrinsic parameters to efficiently evaluate the likelihood first introduced by [22]. Reference that work for more a more detailed description.

1.3.1.1 Waveform Decomposition

The GW strain measured by the kth detector is given by

$$h_k(t) = F_{+,k}(\delta, \alpha, \psi)h_{+,k}(t) + F_{\times,k}(\delta, \alpha, \psi)h_{\times,k}(t) \tag{1.3.26}$$

where $F_{+,k}, F_{\times,k}$ are the antenna patterns of the detector and $h_{+,k}, h_{\times,k}$ are the two components of the gravitational wave strain, evaluated at the kth detector. While the antenna patterns depend only on the extrinsic sky location and polarization angle, the polarizations depend on both intrinsic and extrinsic parameters. At leading order for inspiral-only waveforms, the polarizations are described by the orbital phase and the post-Newtonian "velocity" parameters. These depend only on combinations of the masses of the binary. We can then assume the functional form of a -2 spin-weighted spherical harmonic mode decomposition (denoted h_{lm}) to further separate intrinsic and extrinsic parameters appearing in the polarizations as

$$h_{+,k}(t) - ih_{\times,k}(t) = \frac{D_{\rm ref}}{D}\sum_{lm}\hat{h}_{lm}(\mathcal{M}, \eta, t_k; t)Y^{-2}_{lm}(\iota, -\phi_c) \tag{1.3.27}$$

evaluated at some fixed distance $D_{\rm ref}$. If we define a complex-valued antenna pattern for each detector as

$$F_k = F_{+,k} + iF_{\times,k}, \tag{1.3.28}$$

then we can re-express the measured strain in the kth detector as

$$h_k(\vec{\lambda}, \vec{\theta}; t) = \mathrm{Re}\,\frac{D_{\rm ref}}{D}F_k(\alpha, \delta, \psi)\sum_{lm}\hat{h}_{lm}(\mathcal{M}, \eta, t_k; t)Y^{-2}_{lm}(\iota, -\phi_c). \tag{1.3.29}$$

Besides t_k, we have now fully separated the intrinsic parameter (entering only in $\hat{h}_{lm}$) from the extrinsic parameters (entering only in F_k and Y^{-2}_{lm}).

1.3.1.2 Evaluating the Likelihood

Eq. 1.3.29 shows that we can take the observed strain from a detector and rewrite it into a linear combination of a harmonic mode time series. As stated above, all the intrinsic parameters are contained in $\hat{h}_{lm}$ while all the extrinsic parameters (aside from t_k) are contained in F_k and Y^{-2}_{lm}. As shown in Eq, 1.1.14, the likelihood calculations only involve an inner product. Because of this, we can pull these extrinsic-dependent coefficients outside the inner product integral. We can therefore calculate the inner products involving $\hat{h}_{lm}$ and d once. We can then calculate the likelihood for many different extrinsic parameters by simple re-evaluating the coefficients and reconstructing the linear combination. To that end, we define the following:

$$\begin{aligned} Q_{k,lm}(\vec{\lambda}, t_k) &= \langle h_{lm}(\vec{\lambda}, t_k) | d \rangle_k \\ &= 2 \int_{\infty}^{\infty} \frac{df}{S_k(|f|)} \exp^{2\pi i f t_k} \tilde{h}^*_{lm}(\vec{\lambda}; f) \tilde{d}(f), \end{aligned} \tag{1.3.30}$$

$$U_{h_{lm},l'm'}(\vec{\lambda}) = \langle h_{lm} | h_{l'm'} \rangle_k, \tag{1.3.31}$$

$$V_{h_{lm},l'm'}(\vec{\lambda}) = \langle h^*_{lm} | h_{l'm'} \rangle_k. \tag{1.3.32}$$

If we plug Eq. 1.3.29 into Eq. 1.1.14, take the log and collect terms, we get the following:

$$\begin{aligned} \ln \mathcal{L}(\vec{\lambda}; \vec{\theta}) = (D_{\rm ref}/D) \, \mathrm{Re} \sum_k \sum_{lm} (F_k Y^{(-2)}_{lm})^* Q_{k,lm}(\vec{\lambda}, t_k) \\ - \frac{(D_{\rm ref}/D)^2}{4} \sum_k \sum_{lml'm'} \Big[|F_k|^2 [Y^{(-2)}_{lm}]^* Y^{(-2)}_{l'm'} U_{h_{lm},l'm'}(\vec{\lambda}) + \mathrm{Re} \left(F_k^2 Y^{(-2)}_{lm} Y^{(-2)}_{l'm'} V_{h_{lm},l'm'}(\vec{\lambda}) \right) \Big]. \end{aligned} \tag{1.3.33}$$

There are a couple things to note about Eq. 1.3.33. Firstly, the large computational cost involved in evaluating Eq. 1.3.33 is dominated by the $Q_{k,lm}, U_{h_{lm},l'm'}, V_{h_{lm},l'm'}$. This is where the orbital dynamics, the h_{lm}, the inner products, and the inverse Fourier transforms are all calculated. Secondly, the cost of calculating the F_k and Y^{-2}_{lm} are much cheaper. Thirdly, we can therefore efficiently evaluate the likelihood by calculating the $Q_{k,lm}, U_{h_{lm},l'm'}, V_{h_{lm},l'm'}$ at a fixed intrinsic point and varying the extrinsic parameters. We therefore efficiently integrate over the extrinsic parameters and obtain a marginalized likelihood over (and, by implication,

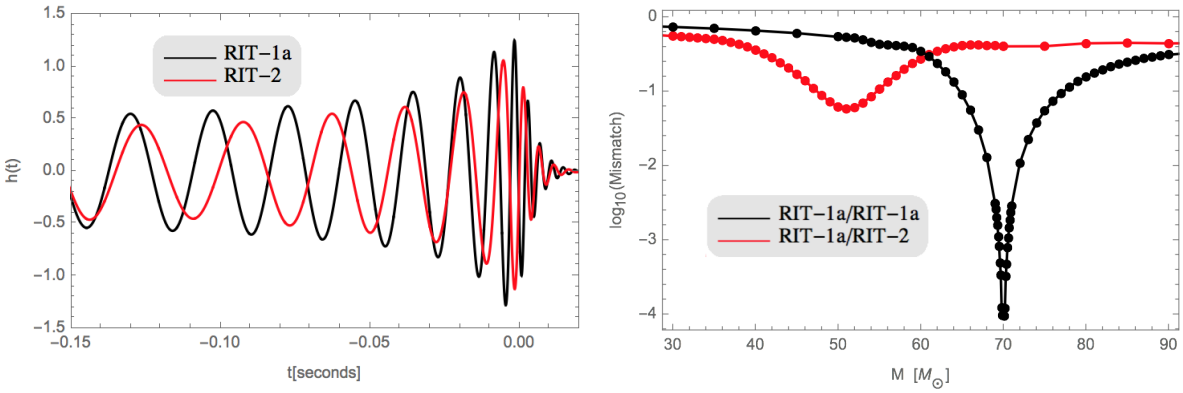

Figure 1.1: **Example - Assessing differences between two NR simulations with different parameters - I**: Two representations of different aligned spin binaries, RIT-1a and RIT-2, with mass ratios $q = 1.22$ and $q = 2.0$ respectively, illustrating how dramatic differences propagate into our diagnostics. *Left panel*: The strain along a line of sight inclined at $\iota = 0.785$ and evaluated for a total mass $M = 70M_{\odot}$. *Right panel*: The mismatch between synthetic data and candidate templates as a function of the template's total mass. The black curve is the null result (mismatch of one simulation with itself). The red curve shows the mismatch between the two simulations.

a posterior for) the intrinsic parameters. Another thing to note is that the computations for each intrinsic point are completely independent from each other making the algorithm highly-parallelizable. For more details, see [22].

1.3.2 Testing and Validating parameter inference using NR waveforms

Because RIFT can be applied to any waveform, including numerical relativity, RIFT can and has been used to directly infer the parameters of GW sources by comparison to numerical relativity simulations alone. This approach was the subject of my master's thesis, which both introduced the approach and validated it [94, 18]. This involved investigating multiple potential sources of systematic error with multiple quantitative diagnostics. One of these diagnostics is called the mismatch. This is calculated by the following equation:

$$\mathcal{M}(h_0, h) = 1 - P(h_0, h) \tag{1.3.34}$$

where h and h_0 are two complex GW strains that you wish to compare and the function P is

$$\ln \mathcal{L}_{\rm single,max} = \max_{\psi, D_L, t_{\rm event}} \ln \mathcal{L}_{\rm single} = -\rho^2(1-P), \tag{1.3.35}$$

$$P(h_0, h) \equiv \max_{\psi, t_{\rm event}} \frac{|\langle h_0|h\rangle|}{\sqrt{\langle h_0|h_0\rangle\langle h|h\rangle}}. \tag{1.3.36}$$

As Equation 1.3.35 implies, this is maximized over polarization angle and event time. Again the $\langle a|b\rangle$ notation represents an inner product calculated by Equation 1.1.9. The mismatch identifies differences between two waveforms. Figure 1.1 shows an example of this type of comparison between two intrinsically different NR simulations (both systems' total mass is $M = 70$ solar masses $\mathrm{M}_\odot$). The left panel shows the GW strain as a function of time. Just by looking at this panel, one can see significant differences between the waveforms. This is further shown in the right panel, which shows the the mismatch as a function of the total mass of the system. The black curve shows the null result: one of the NR simulations compared to itself. The red curve shows the comparison between the two simulations. While the mismatch between the simulation and itself approaches zero as total mass approaches 70 $\mathrm{M}_\odot$, the minimum mismatch when comparing the two simulations is only $\sim 10^{-1.1}$. The total mass at which the minimum mismatch is reach has also shift from the true total mass by ~ 20 $\mathrm{M}_\odot$. This confirms what can already see by eye: that the two waveforms are intrinsically different. This analysis can be applied on any potential error source in question.

Another diagnostic we used is the marginalized likelihood($\ln \mathcal{L}_{\rm marg}(M)$) for a single simulation on some reference data (e.g., the simulation itself, or a signal with comparable physical origin). This function enters naturally into our full parameter estimation calculation; therefore, it allows us to test all of the quantities that influence our principal result directly. Using our example again, the left panel of Figure 1.2 shows the two $\ln \mathcal{L}_{\rm marg}(M)$ curves. The colors here are similar to the colors in Figure 1.1; the black curve represents the comparison between one of the simulations with itself, and the red curve represents the comparison between the two different simulations. Not only do we see this shift in the total mass corresponding to the highest $\ln \mathcal{L}_{\rm marg}$, but we also notice that this $\max \ln \mathcal{L}_{\rm marg}$ is lower than when comparing the

ILE run (source/template)	D_{KL}
Sim1/Sim1	0.0
Sim1/Sim2	288.8

Table 1.1: **KL Divergence between two PDFs**: This table shows two KL Divergence calculations between a PDF and itself as well as the two different PDFs.

simulation to itself. The $\ln \mathcal{L}_{\rm marg}$ can therefore be thought of as a quantity that measures the similarity between the data and the template. To better quantify this result, we converted $\ln \mathcal{L}_{\rm marg}(M)$ into a posterior density function (PDF) by the following equation:

$$p_c(M) = \frac{1}{\int dM e^{\ln \mathcal{L}_{\rm marg}}} e^{\ln \mathcal{L}_{\rm marg}}. \tag{1.3.37}$$

For our example, this is illustrated in the right panel of Figure 1.2. The colors represent the same as in the left panel. The peak mass shift is again visible as in the top panel. To quantify this difference more, we calculate the KL Divergence: a measurement of the similarity of two PDFs. This can be calculated by the following equation:

$$\begin{aligned} D_{KL}(p_*|p) &= \int dx p(x) \ln p(x)/p_*(x) \\ &= \ln \frac{\sigma}{\sigma_*} - \frac{1}{2} + \frac{(\bar{x} - \bar{x}_*)^2 + \sigma_*^2}{2\sigma^2}. \end{aligned} \tag{1.3.38}$$

The KL Divergence of the null test and the two simulations can be found in Table 1.1. In congruent with the previous diagnostics, the high KL Divergence value for the two PDFs suggest that the two distributions are intrinsically different. While this was an extreme case, these same diagnostics were used effectively to investigate potential sources of error. These included the finite duration of NR waveforms, the Monte Carlo error, the extraction of the NR waveforms, and the resolution of the NR waveforms. For more details, see [18].

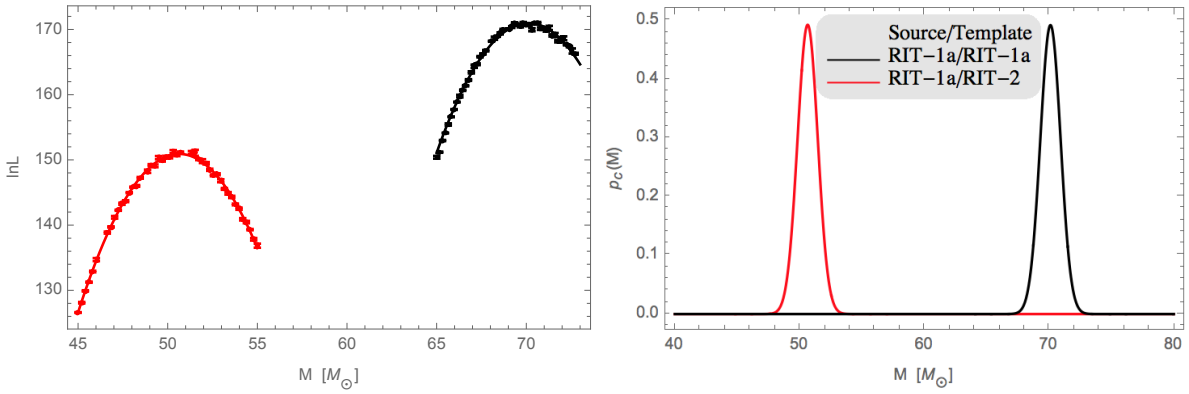

Figure 1.2: **Example - Assessing differences between two NR simulations with different parameters - II**: Two representations of different aligned spin binaries, RIT-1a and RIT-2, with mass ratios $q = 1.22$ and $q = 2.0$ respectively, illustrating how dramatic differences propagate into our diagnostics. *Left panel*: Points show the marginalized likelihood versus total mass calculated by applying the same template simulation to two different sources: itself in black and the different simulation in red. *RIght panel*: The corresponding one-dimensional posteriors $p_c(M)$ [Eq. (1.3.37)]. Both panels illustrate how an ill-suited simulation with large mismatch (i.e., the red curve) correlates with a drastic shift in parameters (here, total mass) relative to the true best-fit solution (here, the black curve). Also, the ill-matched simulation cannot recover all the information available to the true solution, so the peak $\ln \mathcal{L}_{\rm marg}$ for the red curve is substantially lower ($\simeq 20$) than the peak of the black curve.

1.4 Organization

The rest of this dissertation is organized as follows. Chapter 2 is a methods study focusing on establishing and validating a novel PE algorithm. In this chapter, we demonstrate via some anecdotal and systematic examples that the algorithm can infer the parameters of coalescing compact binaries with speed and accuracy. We also include applications that the algorithm is uniquely designed to perform. Chapter 3 is a collection of different studies I contributed to focusing on the effects of HOM as well as waveform systematics between different waveform models described in Section 1.2. In almost all studies, RIFT played a key role. Section 3.2 is a study where we focus on the waveform systematics between two common precessing waveforms and how they can affect are interpretation of real gravitational wave sources. Section 3.3 descriptions a NR follow up study based of the PE done for GW170104. In this project, we demonstrate that the full numerical solution includes information not accurately captured in the semi-analytical models. We quantify these differences by comparing the predictions of

these simulations and the models to the observations of GW170104. Section 3.4 describes a PE study where we evaluate the impact of HOM using a synthetic set of sources with a wide range of parameter values. We then use the same model to recover these parameters multiple times including different number of HOMs. Section 3.5 describes an analysis of the loudest trigger in LIGO's 2015-2017 data set. If real, the analysis from this trigger recovers a total mass consistent with that of an intermediate mass BH. Chapter 4 describes the LVC work analyzing some of the events from the first half of the 3rd Observing Run. Chapter 5 wraps up this dissertation: summarizing the other chapters as well as discussing the impact this work has had on the field. There is also a brief discussion of future projects derived from this work.

1.5 Acronyms

In this section, I summarize the acronyms used in this dissertation.

- RIFT: Rapid parameter inference on gravitational wave sources via Iterative FiTting
- GW(s): gravitational wave(s)
- BBH(s): binary black hole(s)
- LIGO: Laster Interferometer Gravitational Wave Observatory
- LVC: LIGO-Virgo Collaboration
- PE: parameter estimation
- BH(s): black hole(s)
- EOB: effective-one-body
- NR: numerical relativity
- NRSur: NR surrogate
- HOM(s): higher-order mode(s)

- NS(s): neutron star(s)
- BNS(s): binary neutron star(s)
- RIT: Rochester Institute of Technology
- GT: Georgia Tech
- SXS: Simulation eXtreme Spacetimes Collaboration
- SpEC: Spectral Einstein Code
- LI: LALInference
- ILE: integrate likelihood over extrinsic
- CIP: construct intrinsic posterior
- GP: gaussian process
- PSD(s): power spectral density
- PDF: posterior density function
- CDF: cumulative density function
- MaP: maximum a posteriori
- SNR: signal-to-noise
- JSD: Jensen-Shannon divergence
- KL Divergence or KLD: Kullback-Leibler divergence

Chapter 2

Rapid and accurate parameter inference for coalescing, compact binaries

In this chapter, we introduce the parameter inference technique that forms the foundation of my dissertation work: RIFT.

- In Section 2.1, the methodology for each part of the code is explained.
- In Section 2.2, we briefly go over the models used in this chapter.
- Section 2.3 includes validation comparisons to the other standard PE tools.
- Section 2.4 describes a wide range of unique applications for which RIFT has a natural affinity with Section 2.5 describing a reanalysis of GW150914.

The contents of this chapter are taken from [1].

2.1 Parameter inference via iterative gaussian process fits

2.1.1 Coordinates and notation

For a detailed description of the parameters used in this chapter, see Section 1.1.2. For this chapter, we define the mass ratio to be:

$$q = m_2/m_1, \text{where} \quad m_1 > m_2. \tag{2.1.1}$$

2.1.2 Marginalized likelihoods

ILE – a specific algorithm to "integrate over extrinsic parameters" – provides a straightforward and efficient mechanism to compare any specific candidate gravitational wave source with real or synthetic data [22, 9, 18, 95], by marginalizing the likelihood of the data over the seven coordinates characterizing the spacetime coordinates and orientation of the binary relative to the earth. Specifically the likelihood of the data given gaussian noise has the form (up to normalization)

$$\ln \mathcal{L}(\boldsymbol{\lambda}; \theta) = -\frac{1}{2} \sum_k \langle h_k(\boldsymbol{\lambda}, \theta) - d_k | h_k(\boldsymbol{\lambda}, \theta) - d_k \rangle_k - \langle d_k | d_k \rangle_k, \tag{2.1.2}$$

where h_k are the predicted response of the k^{th} detector due to a source with parameters ($\boldsymbol{\lambda}$, θ) and d_k are the detector data in each instrument k; $\boldsymbol{\lambda}$ denotes the combination of redshifted mass M_z and the remaining parameters needed to uniquely specify the binary's dynamics; θ represents the seven extrinsic parameters (4 spacetime coordinates for the coalescence event and 3 Euler angles for the binary's orientation relative to the Earth); and $\langle a|b\rangle_k \equiv \int_{-\infty}^{\infty} 2df \tilde{a}(f)^* \tilde{b}(f)/S_{h,k}(|f|)$ is an inner product implied by the k^{th} detector's noise power spectrum $S_{h,k}(f)$. In practice we adopt a low-frequency cutoff $\mathrm{f_{min}}$ so all inner products are modified to

$$\langle a|b\rangle_k \equiv 2 \int_{|f|>f_{\min}} df \frac{[\tilde{a}(f)]^* \tilde{b}(f)}{S_{h,k}(|f|)}. \tag{2.1.3}$$

The joint posterior probability of $\boldsymbol{\lambda}, \theta$ follows from Bayes' theorem:

$$p_{\rm post}(\boldsymbol{\lambda}, \theta) = \frac{\mathcal{L}(\boldsymbol{\lambda}, \theta) p(\theta) p(\boldsymbol{\lambda})}{\int d\boldsymbol{\lambda} d\theta \mathcal{L}(\boldsymbol{\lambda}, \theta) p(\boldsymbol{\lambda}) p(\theta)}, \tag{2.1.4}$$

where $p(\theta)$ and $p(\boldsymbol{\lambda})$ are priors on the (independent) variables $\theta, \boldsymbol{\lambda}$. For each $\boldsymbol{\lambda}$, we evaluate the marginalized likelihood

$$\mathcal{L}_{\rm marg} \equiv \int \mathcal{L}(\boldsymbol{\lambda}, \theta) p(\theta) d\theta \tag{2.1.5}$$

via direct Monte Carlo integration, where $p(\theta)$ is uniform in 4-volume and source orientation. To evaluate the likelihood in regions of high importance, we use an adaptive Monte Carlo as described in [22]. As described in Pankow et al, this marginalized likelihood can be evaluated efficiently because, having generated the dynamics and outgoing radiation in all possible directions relative to the binary once and for all for fixed λ, the likelihood can be evaluated as a function of θ at very low computational cost.

Using Bayes' theorem, the posterior distribution for intrinsic parameters $\boldsymbol{\lambda}$ can be expressed as

$$p_{\rm post} = \frac{\mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})}{\int d\boldsymbol{\lambda} \mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})}. \tag{2.1.6}$$

where prior $p(\boldsymbol{\lambda})$ is the prior on intrinsic parameters like mass and spin. ILE itself will only provide point estimates $\ln \mathcal{L}_{{\rm marg}\alpha}$ given proposed evaluation points $\boldsymbol{\lambda}_\alpha$, where α indexes the points supplied to ILE, not the interpolated $\ln \mathcal{L}_{\rm marg}$ necessary to construct a full solution.

2.1.3 Gaussian process interpolation

Given some proposed training data $\{(\boldsymbol{\lambda}_\alpha, \ln \mathcal{L}_{{\rm marg}\alpha})\}$, we estimate $\ln \mathcal{L}_{\rm marg}$ via conventional Gaussian Process (GP) interpolation [96] to produce a weakly nonparametric interpolation and error estimate. Gaussian process interpolation has already been used in the field [97, 98, 95, 99, 100, 95], particularly to propagate uncertainties. [For brevity and to be consistent with conventional notation, in this section we denote $\boldsymbol{\lambda}_\alpha$ by x and $\ln \mathcal{L}_{{\rm marg}\alpha}$ by y.] In this approach,

we estimate the expected value of $y(x)$ from data x_* and values y_* via

$$\langle y(x)\rangle = \sum_{\alpha,\alpha'} k(x, x_{*,\alpha})(K^{-1})_{\alpha,\alpha'} y_{*,\alpha'} \tag{2.1.7}$$

where α is an integer running over the number of training samples in (x_*, y_*) and where the matrix $K = k(x_\alpha, x'_\alpha)$ y_*. We employ a kernel function $k(x, x')$ which allows for uncertainty in each estimated training point's value $y_{*,\alpha}$ due to Monte Carlo integration, as well as a conventional squared exponential kernel to allow for changes in the functions versus parameters:

$$k(x, x') = \sigma_o^2 e^{-(x-x')Q(x-x')/2} + \sigma_n^2 \delta_{x,x'} \tag{2.1.8}$$

Following usual practice and to insure flexibility for generic sources, we optimize the hyperparameters σ_o, σ_n and the positive-definite symmetric matrix Q on our training data. We perform all gaussian process interpolation with widely-available open-source software [101]. The computational cost of full-scale Gaussian process optimization and evaluation increases rapidly with the dimension D of the matrix K, as D^3 and D^2 respectively; for typical hardware, we are presently limited to $O(10^4)$ training points in 8 dimensions. In future work we will employ other techniques like sparse approximations and multilayer ("pool-of-experts") designs, which can achieve comparable results with better scaling.

2.1.4 Flexible Monte Carlo generation of posterior samples

We use a single code (henceforth denoted by CIP) to Construct the Intrinsic Posterior distribution, by loading training data $\{(\boldsymbol{\lambda}_\alpha, \ln \mathcal{L}_{\mathrm{marg}\alpha})\}$; fitting that data, using some coordinate system X for $\{\boldsymbol{\lambda}\}$; and then constructing samples from the posterior distribution via adaptive Monte Carlo using another coordinate system Y with fiducial prior distributions $p(y)$, employing the same adaptive Monte Carlo techniques described [22] and applied in ILE. When fitting the likelihood, we employ coordinate systems well-adapted to the likelihood, which based on many decades of theoretical and computational studies are likely to produce an approximately gaussian likelihood in the limit of strong signals [e.g., coordinates aligned with a quadratic ap-

proximation to the log-likelihood in the high- signal-to-noise (SNR) limit]. When performing a Monte Carlo integral to construct a posterior distribution, however, we employ coordinates Y that allow us to efficiently specify the priors in separable form and simultaneously sample the parameter space thoroughly in the region with significant prior support. For example, we often use the chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ as a sampling coordinate, to insure our adaptive Monte Carlo method can efficiently identify the often exceptionally narrow region of $\mathcal{M}$ consistent with the posterior, particularly for binary neutron stars. In Appendix 2.6 we describe the specific array of coordinate systems and priors we use.

As with ILE, the adaptive Monte Carlo produces a sequence of weighted points $w_k, \boldsymbol{\lambda}_k$ that fully characterize the posterior distribution, such that the integral of any function $G(\boldsymbol{\lambda})$ with support inside the posterior can be evaluated by Monte Carlo as $[\sum_k G(\boldsymbol{\lambda}_k) w_k]/\sum_q w_q$. Specifically, relative to some reference prior $p_{\rm ref}(\boldsymbol{\lambda})$ identified during adaptive Monte Carlo integration, the samples λ_k are drawn from $p_{\rm ref}$ and the weights are $w_k = \mathcal{L}_{\rm marg}(\boldsymbol{\lambda}_k) p(\boldsymbol{\lambda}_k)/p_{\rm ref}(\boldsymbol{\lambda}_k)$. For example, any one-dimensional marginal distribution for a function $g(\boldsymbol{\lambda})$ can be computed via the corresponding weighted sum $P(< G) = [\sum_k w_k \Theta(g(\boldsymbol{\lambda}_k) - G)]/\sum_q w_q$, where Θ is the Heavyside function. To improve compatibility with other codes' output and to minimize overhead – in practice, a set of weighted posterior samples often includes $\gg 10^8$ points, many with low weight – we then uniformly resample from the weighted posterior points, with replacement. Specifically, after ordering the sequence w_k so $w_1 \leq w_2 \leq \ldots$, we draw N uniformly distributed random numbers u_q and choose $\boldsymbol{\lambda}_q$ such that q is the largest number with $\sum_{k \leq q} w_k/\sum_k w_k < u_q$.

Our adaptive Monte Carlo procedure currently requires separable priors, as described in Appendix 2.6. We generate results for generic priors by reweighting the underlying weighted samples produced above, before performing the draw-with-replacement procedure to generate a fair sample from the target posterior. For example, if $\{(w_k, \boldsymbol{\lambda}_k)\}$ are generated by the procedure above with a prior $p_{\rm ref}(\boldsymbol{\lambda})$, then we generate posterior distributions suitable to a prior $p(\boldsymbol{\lambda})$ via reweighted samples $\{(w_k p(\boldsymbol{\lambda}_k)/p_{\rm ref}(\boldsymbol{\lambda}_k), \boldsymbol{\lambda}_k)\}$, following the procedures outlined above.

2.1.5 Iterative procedure

The fitting and sampling procedures described above produce a proposed set of posterior samples $B_0 = \{\boldsymbol{\lambda}_k\}$, given training data $A_0 = \{\boldsymbol{\lambda}_\alpha\}$. We then use ILE to evaluate $\ln \mathcal{L}_{\rm marg}$ on $\boldsymbol{\lambda}_k$, then perform the CIP procedure described in Section 2.1.3 (gaussian process interpolation) and 2.1.4 (posterior generation) starting with revised training data $A_1 = A_0 \cup B_0$ to produce a new set of proposed samples B_1. By repeating this process several times, we can validate and refine our fit and hence posterior. To assess whether the fit and posterior have converged, we use (all of) the one-dimensional marginal distributions, comparing the empirical cumulative distributions in one iteration against the next using standard tools; see Appendix 2.7. As described in Appendix 2.7, we tested this iterative fitting and posterior generation procedure using synthetic likelihood functions $\ln \mathcal{L}_{\rm marg}$ for which the posterior distributions can be calculated analytically. This procedure succeeds consistently for all dimensions we thoroughly tested ($d \leq 6$).

We extended the basic framework introduced above to make it more robust and efficient. For example, the procedure above can break down if the training data doesn't contain all of the posterior (e.g., if the training data has a hard cutoff that the posterior doesn't). To make the process more robust against this possibility, we employ a dithering step. Specifically, given (a subset of) B_k and its covariance Σ in the subset of parameters we choose to dither, we can add random offsets $\Delta\boldsymbol{\lambda}_k$ drawn from a multinormal distribution with covariance $\epsilon^2\Sigma$, where ϵ is a factor controlling the amount of dithering. After removing samples which due to $\Delta\boldsymbol{\lambda}_k$ are no longer physical, we have a new set C_k. The iterative procedure uses both undithered and dithered points, so $A'_{k+1} = A'_k \cup B_k \cup C_k$. We typically employ a dithering factor $\epsilon \simeq 1$. Second, to make the process more efficient, we can (if needed) perform the first several iterations with fewer and physically-motivated degrees of freedom: the map $X(\boldsymbol{\lambda})$ used to perform the fit will have fewer dimensions than $\boldsymbol{\lambda}$. This approach may be required if the initial training set A_0 is too small to provide a useful fit in all dimensions of $\boldsymbol{\lambda}$. If the likelihood can be approximated at some level using only the variables in X – if only that subset of variables characterizes the most easily measurable quantities – then these dimensionally-reduced iterations cause the

remaining degrees of freedom to be efficiently and randomly explored. As a concrete example, for massive BBHs like GW150914, even if precessing, the likelihood can be well approximated using just $\mathcal{M}, \eta, \chi_{\rm eff}$ [9], though after more iterations our fitting procedure captures more subtle features of the fit and hence posterior. After these first few seed iterations, we revert back to the conventional approach described above.

Gaussian processes are expensive, with evaluation cost scaling as the number of points squared. Before fitting, we always eliminated points whose marginalized likelihoods were many orders of magnitude outside the expected support of the distribution (i.e., with $\ln \mathcal{L} \lesssim \ln \mathcal{L}_{\rm max} - F\chi^2_d(0.9)/2$ where $\chi^2_d(x)$ is the inverse-chisquared distribution with d degrees of freedom, and $F \simeq 10$). In rare cases, even after this condition is applied, we iteratively accumulate a number of points N greater than our computationally-tractable limit N_*. In these cases, we randomly subdivide our training points into G equal-sized subgroups smaller than N_*. In one approach, we repeat our analysis on each random subsample, constructing $b = 1 \ldots G$ gaussian-process approximations $g_b(\boldsymbol{\lambda})$ and posterior samples A_b, requiring consistency between these outputs A_b. In another approach, we construct the posterior distribution using the average $\bar{g}(\boldsymbol{\lambda}) \equiv \sum_b g(\boldsymbol{\lambda})/G$.

As in previous work [22, 9], we analyze data containing candidate signals starting with a good approximation to (some of) that sources' parameters, provided by the search algorithm which flagged this stretch of data as a candidate binary coalescence. We use this information to conservatively identify the initial grid A_0 of test parameters to explore; see, e.g., [22] and references therein.

2.1.6 Reconstructing source-frame binary masses

ILE compares waveforms with fixed detector-frame masses $m_{i,z} = m_i(1+z)$ to real and synthetic observational data. As a result, the procedures described above (and in all previous work) produce posterior distributions as a function of $M_z = (m_1+m_2)(1+z)$ and dimensionless intrinsic variables. We must perform some additional post-processing and analysis to recover the distribution of the total source-frame mass $M = (m_1+m_2)$ and its joint distribution with

all other intrinsic parameters.

The most straightforward and robust procedure to reconstruct the full, joint source-frame posterior reprocesses the output of ILE from the final set of draws A_N from the iterative procedure described above. Each point $\boldsymbol{\lambda}_k \in A_N$ is a fair draw from the posterior distribution, by design. In the process of performing the Monte Carlo integral for $\ln \mathcal{L}_{\rm marg}$ for each $\boldsymbol{\lambda}_k$, ILE produces weighted posterior samples $\{w_{k,\alpha}, \theta_{k,\alpha}\}$, where α indexes the extrinsic $>$ posterior samples produced for each k, and where the parameters and where the parameters $\theta_{k,\alpha}$ include the distance and therefore redshift z. We generate the joint posterior distribution on source-frame parameters by combining all ILE output samples and re-expressing all masses using a suitable redshift. For example, we can generate the cumulative distribution of M via $P(< M) = \sum_{k,\alpha} w_{k,\alpha} \Theta(M - M_{z,k}/(1+z_{k,\alpha}))$.

2.1.7 Cost and speed comparison versus LALInference

The overall computational and wallclock time of a RIFT analysis depends on the total number $N_{\rm eval}$ of $\ln \mathcal{L}_{\rm marg}$ evaluations needed; the runtime time $\tau_{\rm eval}$ per evaluation, on a single core; the number of nodes $N_{\rm clus}$ available in a given computational cluster for this analysis, presumed $N_{\rm clus} < N_{\rm eval}$; the number of iterations $N_{\rm it}$ needed; and the worst-case runtime $\tau_{\rm fit}$ for the fit-plus-posterior code (CIP). The overall number of core-hours will be $N_{eval}\tau_{eval} + N_{\rm it}\tau_{\rm fit}$; the wallclock run time will be $N_{\rm it}\tau_{\rm fit} + \tau_{\rm eval}N_{\rm eval}/N_{\rm clus}$.

The cost and wallclock run time for RIFT continues to evolve substantially due to ongoing backend improvements. For example, a recent GPU-optimized version of RIFT code performs likelihood evaluations in tens of seconds [102]. [The version employed by the LVC for published work in O2 [103, 15] used an older code version, where typical single-job runtimes had $\tau_{\rm eval} =$ 15 to 30 minutes, depending on the model and accuracy goal, on clusters with typically $N_{\rm clus} = 10^3$ cores available for an analysis.] For the original GP implementation, the GP fitting code could require as long as $\tau_{\rm fit} \simeq 1$ day $\times (N_{\rm eval}/10^4)^3$, which combined with GP memory footprints effectively introduces a limit of $\simeq 10^4$ evaluation points for the current GP implementation. Most of this cost involves optimizing the likelihood versus hyperparameters; once a fit is

produced and stored, we can rapidly and in parallel produce as many posterior samples as needed. Because of the rapid increase in cost above $N_{\rm eval} \simeq 10^4$, we always cap the number of points used in a fit to be smaller than 10^4, drawing those points at random from the set of points with sufficiently high marginalized likelihood. Production-quality analysis of high-dimensional models including tides or precession typically requires $N_{\rm eval} \simeq 3 \times 10^4$ marginalized likelihood evaluations, performed via 5 to 10 iterations as needed depending on the novelty of the system. Therefore, the computational cost for first-time calculations using the GPU version is usually below $N_{\rm eval}/120 \lesssim 250$ core-hours, but the wallclock time is bounded above by $N_{\rm fit}\tau_{\rm fit} \simeq 5$ days, dominated by GP fitting of the last few iterations. [Followup calculations or familiar classes events generally require fewer iterations, as they can leverage previously-explored grids or expertise with qualitatively similar posteriors, shifted to align with the current analysis. To be maximally conservative, our discussion below adopts this worst-case scenario.] During these 5 days, RIFT and ILE have evaluated the marginalized likelihood $N_{\rm eval}$ times. In our current configuration, each marginalized likelihood is performed via a Monte Carlo integral with up to 2×10^6 underlying likelihood evaluations, or $2N_{\rm eval} \times 10^6$ likelihood evaluations overall. This computational and wallclock cost can be usefully compared with a LALInference (LI) [23] analysis using SEOBNRv3 comparable to examples presented later. To provide an optimistic example, after 10 days using 50 independent chains, a LI analysis performed 6×10^7 low-temperature likelihood evaluations and expended roughly 14,000 core-hours to produce roughly 6×10^3 samples, or roughly 60 times the overall computational cost to produce a comparable-quality sample size to the worst-case RIFT analysis above. A more typical example expended 10^5 core-hours to produce only 10^3 samples. Our attempts to employ LI with SEOBNRv3 usually do not terminate cleanly with clear indications of convergence, so these estimates likely understate the true cost needed for robust results. While recent optimized versions of SEOBNRv3 are known to reduce its computational cost substantially [104], and an analytic path forward exists for algorithmic changes to enable even more substantial speed improveents for similar EOB models [105], RIFT likewise continues to undergo algorithmic improvements which reduce its overall CPU and wallclock cost by orders of magnitude [102], and can be

applied to any compact binary waveform approximation at any stage of development.

2.2 Models and sources in this work

In this work, we perform parameter inference with several standard approximations to the outgoing radiation of coalescing binaries, which fall in three families: EOB models [39, 106], specifically SEOBNRv3, SEOBNRv4, and SEOBNRv4_ROM; phenomenological frequency-domain inspiral and merger models, specifically using IMRPhenomD and IMRPhenomPv2 [46]; and surrogate waveforms, directly interpolated from numerical relativity [55, 58, 36], specifically using NRSur7dq2. See Section 1.2 for more details.

For most models, we employ `lalsimulation` implementations of these two approximations [107], provided and maintained by their authors in the same form as used in LIGO's O1 and O2 investigations. For NR surrogate waveforms, we employ the software provided by the authors [108]. We use the implementation of TEOBResumS provided by the authors [109], restricting to $\ell = 2$ modes for the outgoing radiation.

Rather than rely on one of these diverse approximations for our synthetic fiducial sources, we instead almost always use the output of detailed simulations of Einstein's equations, provided by the RIT and SXS group [76, 93, 74]. Table 2.1 provides the properties of our synthetic sources. All of the synthetic sources have parameters qualitatively consistent with the observed BBH population.

In our figure legends, we use a few abbreviations for these approximation names: IMRPhenomD will be denoted IMRD; IMRPhenomPv2 will be denoted IMRP; and NRSur7dq2 will be denoted NRSur.

2.3 Validation examples

In this section, we use both LI and RIFT to infer the parameters of three synthetic sources provided in Table 2.1, using the models described in Section 2.2. First and foremost, we use these examples to demonstrate both LI and RIFT produce equivalent results when used to

ID	Model/Numerical Relativity	q	M ($M_\odot$)	χ_{1x}	χ_{1y}	χ_{1z}	χ_{2x}	χ_{2y}	χ_{2z}	λ_1	λ_2
RIT-1	RIT:D12_q1.00_a-0.25_-0.25_n100	1.00	80.0	0.0	0.0	0.250	0.0	0.0	-0.250	-	-
SEOB-1	SEOBNRv3	0.814	110.0	0.09	0.06	0.32	0.27	0.21	-0.58	-	-
RIT-2	RIT:De10_D10.97_q1.3333_a-0.6_-0.8_n100	0.750	80.0	0.0	0.0	-0.600	0.0	0.0	-0.800	-	-
Tides-1	SEOBNRv4T	0.900	2.88	0.0	0.0	0.0	0.0	0.0	0.0	207	409

Table 2.1: **Parameters of synthetic sources**: This table shows the parameters of all the synthetic sources (waveform approximant models and numerical relativity) used in this paper. q is the mass ratio defined with $q < 1$ (see Eq. 2.1.1), M is the detector-frame total mass, and χ_* are the components of the normalized spins (see Eq. 1.1.23). All luminosity distances are set such that the signal-to-noise ratio is around 20 (SNR$\sim$20). Other extrinsic parameters are the following: inclination angle from the line-of-sight is $\iota = \pi/4$, right ascension is RA=0.57, declination is DEC=0.1, and the polarization angle is $\psi = \pi/4$

analyze the same synthetic data with the same underlying model. To compactly characterize both the analysis method and model, we will use a prefix of LI or RIFT and a postfix of the model name; for example, a LI_SEOBNRv4_ROM analysis was performed with LI using the nonprecessing model SEOBNRv4_ROM. We also use these examples as an opportunity to further demonstrate how systematic differences between models can produce moderately different conclusions about each binary's properties. In that context, we emphasize the contrast with other models and NRSur7dq2, which contains higher modes. As with previous analyses with numerical relativity simulations [9] and keeping in mind finite model duration, we adopt a minimum frequency of 30Hz for NRSur7dq2; however, we adopt a minimum frequency of 20Hz for our other analyses.

Each of the synthetic sources used in this work are created with the "zero noise realization", such that the synthetic detector data is exactly equal to the expected response due to our synthetic source. For our BBH investigations, we assume gaussian detector noise in a two-detector network, characterized by the same noise strain power spectrum adopted in our analysis of GW150914 in this and prior work [9]. For our binary neutron star investigation, we assume gaussian noise in a three-detector network (Hanford/Livingston/Virgo) operating at design sensitivity. Each source has been scaled to have an SNR ~ 20.

2.3.1 Aligned NR Source

The first synthetic source corresponds to the first row of Table 2.1 (RIT-1): an equal-mass BBH with $\chi_{1,z} = 1/4, \chi_{2,z} = -1/4$ and $M = 80M_\odot$. For this analysis, we assume both spins

are parallel to the orbital angular momentum, and adopt a uniform prior distribution in $\chi_{i,z}$ from the maximum to minimum allowed by each approximation. Using RIFT, we estimate the parameters with the following nonprecessing models: SEOBNRv4_ROM, IMRPhenomD, and NRSur7dq2 with the latter using modes with $\ell \leq 3$ and restricting the spins to the orbital angular momentum. Using LI, we estimate the parameters with the same two former models SEOBNRv4_ROM and IMRPhenomD. Figure 2.1 shows our inferred 90% credible intervals for several parameters for each of these algorithms and models, as well as associated one-dimensional marginal posterior distributions.

First and foremost, Figure 2.1 demonstrates that LI and RIFT produce comparable results when analyzing this data with SEOBNRv4_ROM (LI is dotted black, RIFT is solid black) as well as with IMRPhenomD (LI is dotted blue, RIFT is solid blue). Second, due to the well-known good agreement between the nonprecessing models SEOBNRv4_ROM and IMRPhenomD for the (2,2) mode [110, 38], we find very consistent posterior distributions for SEOBNRv4_ROM (black) and IMRPhenomD (blue). Finally and also as expected, because these two models omit higher-order modes, an analysis with an NR surrogate model which includes several higher-order modes (here, $\ell_{\rm max} = 3$) draws sharper conclusions about at least one of the parameters: the binary mass ratio. [Conclusions about the other parameters are not significantly impacted by including higher-order modes.] A similar result was found with a reanalysis of GW150914 where the LVC compared the data directly to NR [9].

2.3.2 Precessing NR Source

To test our approach on a precessing source, we generated a synthetic precessing signal using SEOB-1 from Table 2.1, corresponding to a comparable-mass system with modest spins perpendicular to the orbital plane. We adopt a volumetric prior on each BH's spin; see Appendix 2.6.2. With as similar settings as possible, we again estimate the parameters with RIFT and LI but this time using the precessing frequency-domain IMRPhenomPv2 model. Figures 2.2 shows our inferred 90% credible interval for the model for each PE technique, again with LI represented with dotted lines and RIFT results with solid lines. As expected, the results from

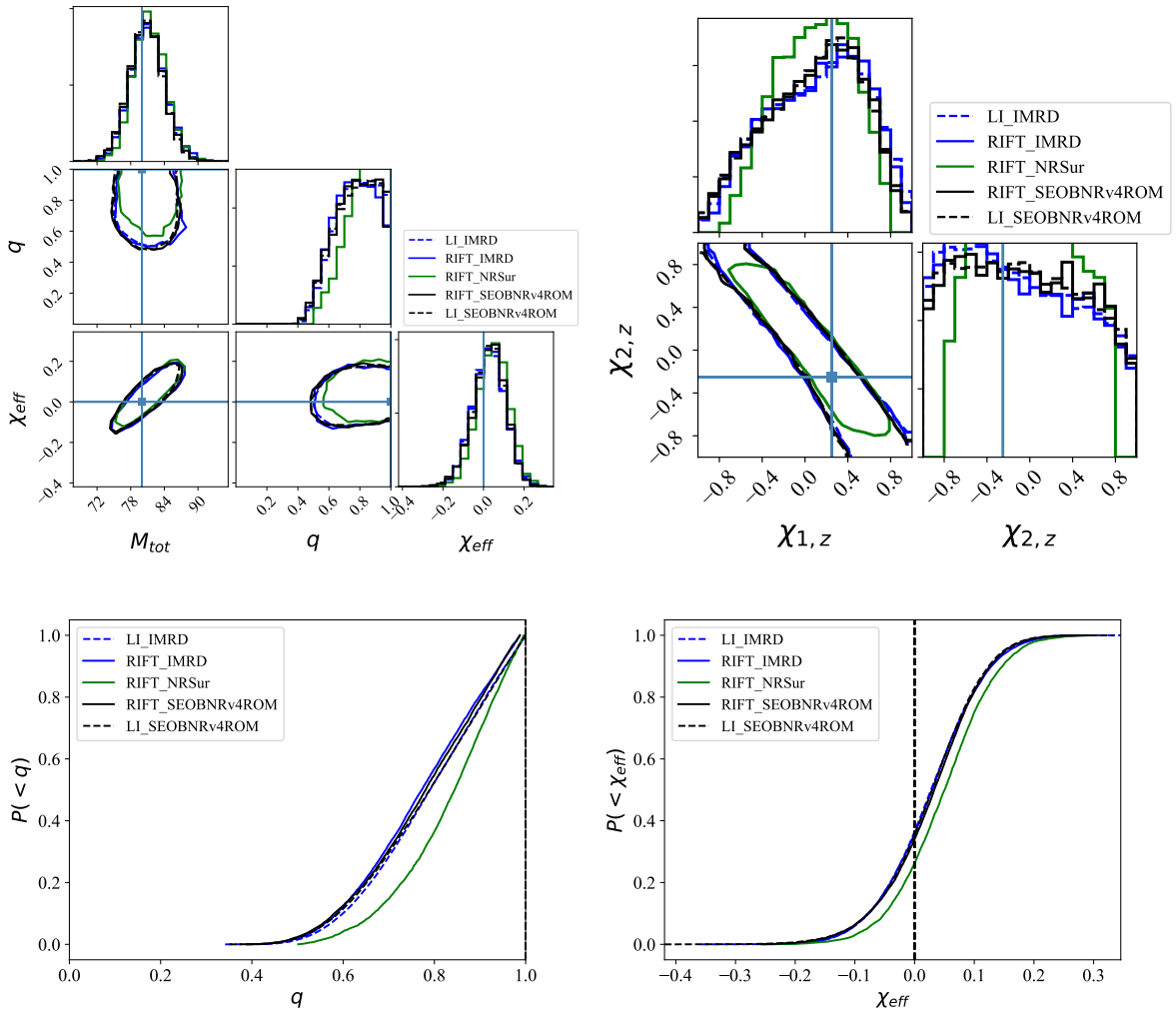

Figure 2.1: **Recovery of nonprecessing NR source with nonprecessing analyses**. This figure shows inferred posterior parameter distributions for RIT-1 in Table 2.1, generated using the CORNER package [8]. The top-left panel shows the one- and two-dimensional marginal distributions for $M, q, \chi_{\rm eff}$; the top-right panel shows the one- and two-dimensional marginal distributions for $\chi_{1,z}, \chi_{2,z}$. The bottom two panels show the cumulative distributions of parameters q and $\chi_{\rm eff}$. In the two-dimensional plots, dotted contours are the 90% credible intervals for the LI results using the IMRPHENOMD and SEOBNRV4_ROM models, while the three solid curves show RIFT results produced with IMRPHENOMD (blue, solid), SEOBNRV4_ROM (black, solid) and NRSur7dq2, restricted to aligned spins and including all modes up to $\ell = 3$ (green, solid). Both LI and RIFT produce comparable results when using IMRPHENOMD or SEOBNRV4_ROM, with statistical differences far smaller than model systematic effects (illustrated here with NRSur7dq2).

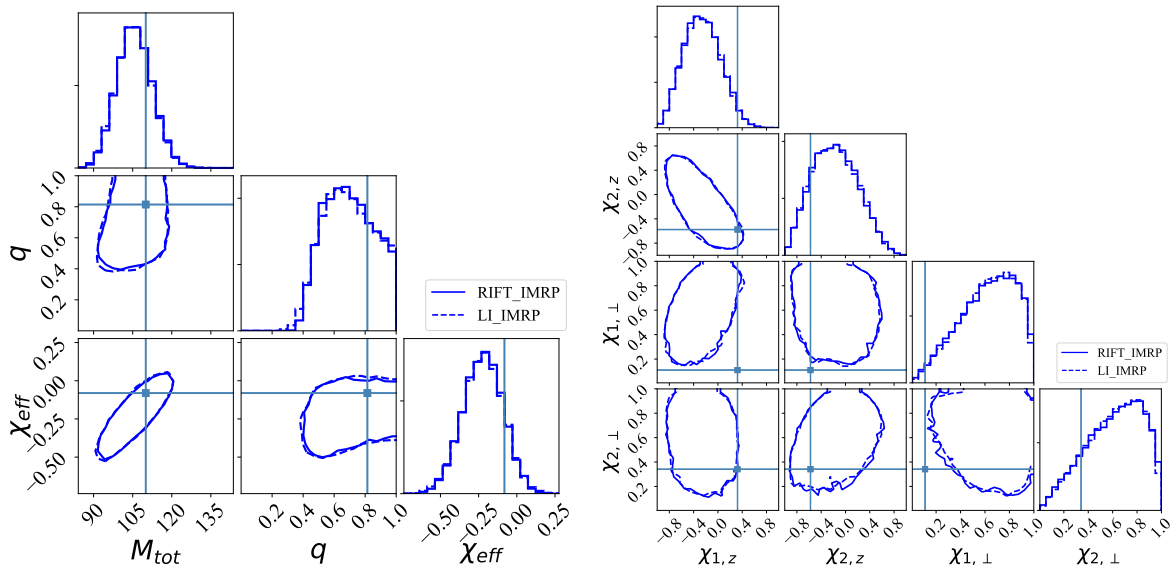

Figure 2.2: **Recovery of a precessing NR source with a precessing model**. This figure shows the results using source SEOB-1 in Table 2.1. The first group of plots shows the PE results in the total mass ($M_{\rm tot}$)-effective spin ($\chi_{\rm eff}$)-mass ratio (q) parameter space (see Eq. 1.1.22 and Eq. 2.1.1 respectively). The second group of plots shows inferences about the various spin components $\chi_{i,z}$, where the z axis convention is defined as parallel to the orbital angular momentum.

the two different algorithms yield consistent results.

Looking at Figure 2.2 closely, one will notice that the true parameters, represented by light blue lines/dots, lie outside some of the 90% CI. This could be due to known differences between IMRPhenomPv2 (the template model) and SEOBNRv3 (the source model), for example as demonstrated in [2]. While the results do not always recover the true parameters of the source, it is encouraging that the two PE techniques are consistent with each other.

2.3.3 Binary neutron star

Due to their low mass, binary neutron stars produce exceptionally long inspirals in the sensitive band of LIGO and Virgo. RIFT can not only efficiently analyze these signals, but do so while using exceptionally costly source models like SEOBNRv4T and TEOBResumS, which can require up to an hour to generate a signal. As a concrete example, Figure 2.3 presents an analysis of synthetic data based on the Tidal-1 entry in Table 2.1. Firstly, the dotted

and solid blue contours show 90% credible intervals and one- and two-dimensional marginal distributions inferred from with LI and RIFT respectively, both using the IMRPhenomD waveform model, modified to include tides [63] (also called IMRPhenomD_NRTidal); both are consistent. Secondly, RIFT analyzed the data with the computationally expensive waveform models: SEOBNRv4T in red and TEOBResumS in orange. These again agree nicely with the previously mentioned IMRPhenomD_NRTidal results as well as the LI-TaylorF2 in green. Using RIFT, we can not only recover the correct posterior, but we can do so with more costly waveforms when analyzing a binary neutron star system.

2.4 Applications

2.4.1 Parameter inference via mixed models

Many approximations are well-suited only to certain parts of the parameter space, breaking down for sufficiently extreme spins, mass ratio, or total mass. For example, NRSur7dq2 is only suitable for mass ratios $q > 0.5$ and $|\chi_i| < 0.8$. RIFT provides an almost-trivial mechanism to flexibly explore parameter inferences that employ different approximations A, B in different regions. Because the computational cost is dominated by ILE, once parameter inference is performed using models A and B separately, we can reanalyze the data with different mixtures of A and B with no additional overhead. For simplicity and to illustrate the method, we will employ the most extreme and simple form of this approach, where the training pairs $(\boldsymbol{\lambda}_k, \ln \mathcal{L}_{\rm marg}(\boldsymbol{\lambda})_k)$ used to produce a fit for $\ln \mathcal{L}_{\rm marg}(\boldsymbol{\lambda})$ are provided using model A when $\boldsymbol{\lambda}_k$ is in some region $\mathcal{V}$, but with model B everywhere outside $\mathcal{V}$.

As a concrete and practical motivation for this strategy, Figures 2.1 and 2.5 (RIT-1 and RIT-2 from Table 2.1 respectively) show analysis of synthetic NR sources with IMRPhenomD and with NRSur7dq2, assuming a nonprecessing binary and adopting a uniform-$\chi_{i,z}$ prior on the two BH spins. The hard limits on NRSur7dq2 both in mass ratio and particularly in $\chi_{i,z}$ have a substantial impact on our conclusions about the binaries parameters especially for systems with large $|\chi_{\rm eff}|$. Because constraints on $\chi_{\rm eff}$ strongly correlate with information

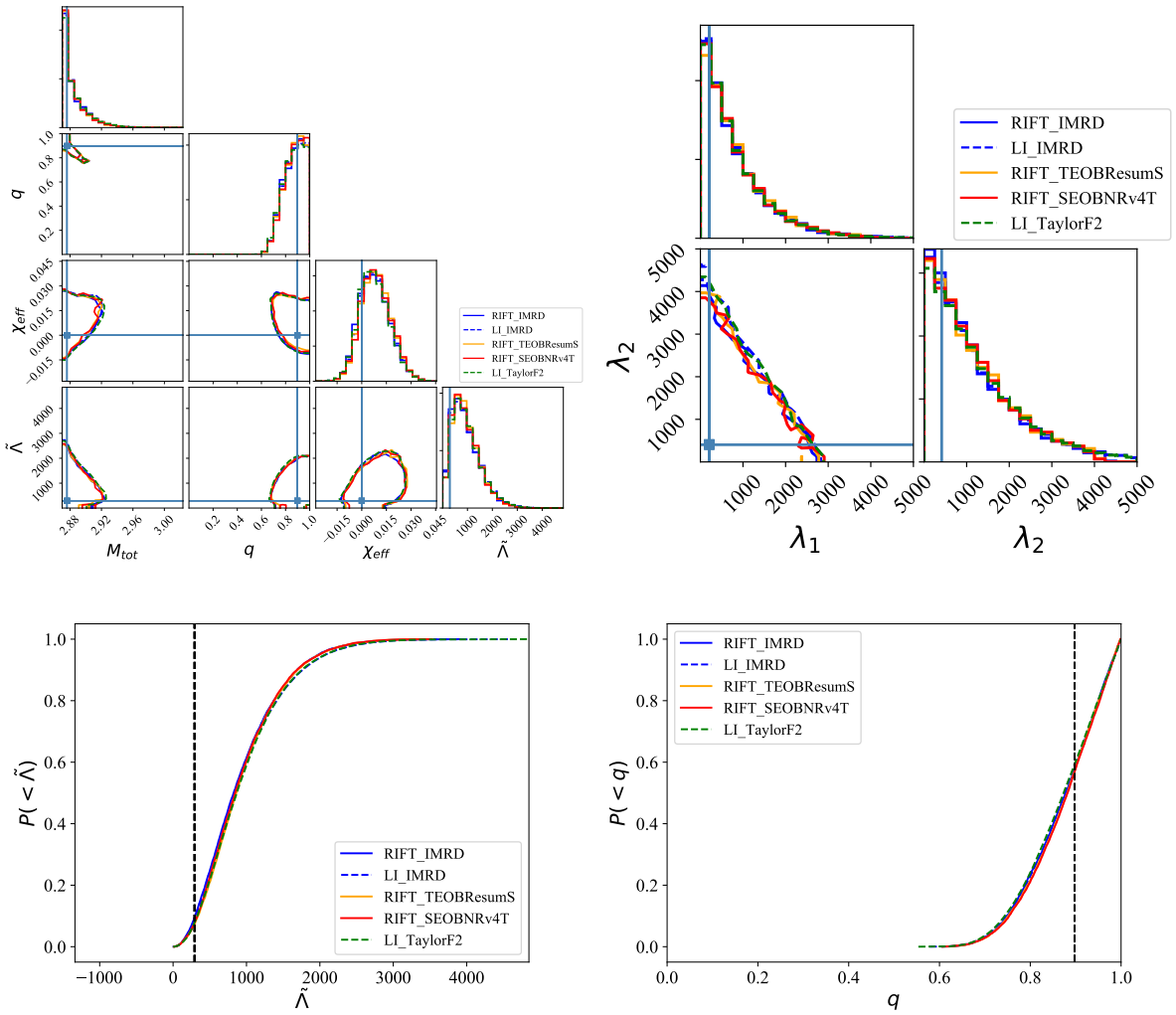

Figure 2.3: **Using sophisticated tidal waveforms to analyze tidal injections** This figure shows inferences about the Tidal-1 source Table 2.1, performed adopting the spin prior in Eq. (2.6.16) with $\chi_{\max} = 0.05$. As in previous figures, the one- and two-dimensional results show all possible one-dimensional posterior distributions and two-dimensional 90% credible intervals for the total mass (M), mass ratio (q), effective tidal deformability ($\tilde{\Lambda}$), and net effective spin ($\chi_{\rm eff}$). As previously, dotted contours and curves correspond to LI results; solid curves are produced with RIFT; and different colors correspond to different waveform models: TaylorF2 is green, IMRPhenomD_NRTidal is blue, SEOBNRv4T is red, and TEOBResumS is orange.

extracted about binary masses, this limitation is not academic, and can complicate attempts to apply these kinds of parameter-limited waveforms to probe effects like waveform systematics.

Figure 2.4 shows an analysis of the same source (RIT-2) as Figure 2.5, using RIFT with a mixed analysis that uses NRSur7dq2 with $\ell \leq 2$ modes in its domain of validity and IMRPhenomD elsewhere. As demonstrated by Figure 2.5, restricting parameter inference to the domain of validity of NRSur7dq2 has a significant impact on the inferred source parameters. By contrast, the mixed model result (represented by the red curve in Figure 2.4) agrees with the full inferences adopted with IMRPhenomD (represented by the blue curve in Figure 2.4). For the purposes of illustration we have intentionally selected a source both where the domain of validity of NRSur7dq2 impacts multiple astrophysical inferences and where extending the model produces comparable results. However, as also illustrated with Figure 2.5, even in this extremely astrophysically pertinent example, these mixed models are being applied in a regime where IMRPhenomD and SEOBNRv4_ROM disagree. In a subsequent investigation, we will follow up this proof- of-concept example with a more detailed analysis of the advantages and applications of this mixed-model strategy, particularly for assessing waveform systematics.

2.4.2 Investigating systematics

The method provides a particularly straightforward but powerful method to assess the impact of systematics on posterior distributions. Simply put, if A and B are two models, we can use precisely the same evaluation points $\boldsymbol{\lambda}_\alpha$ to estimate the marginalized likelihoods according to both models ($\ln \mathcal{L}_\alpha(A)$ and $\ln \mathcal{L}_{\mathrm{marg}\alpha}(B)$), and hence to estimate the corresponding posterior distributions. In practice, we first generate a posterior distribution using the iterative procedure described above for model A (e.g., SEOBNRv4) and then, using all the test points proposed during the iterative scheme, derive the corresponding posterior for model B. This approach provides not only the posterior distributions but also the ingredients needed for a detailed investigation into the origin of any discrepancies: the point-by-point differences between $\ln \mathcal{L}_{\mathrm{marg}\alpha}(A)$ and $\ln \mathcal{L}_{\mathrm{marg}\alpha}(B)$, as a function of model parameters. Once differences are identified, the ability to quickly produce a single scalar diagnostic for model differences

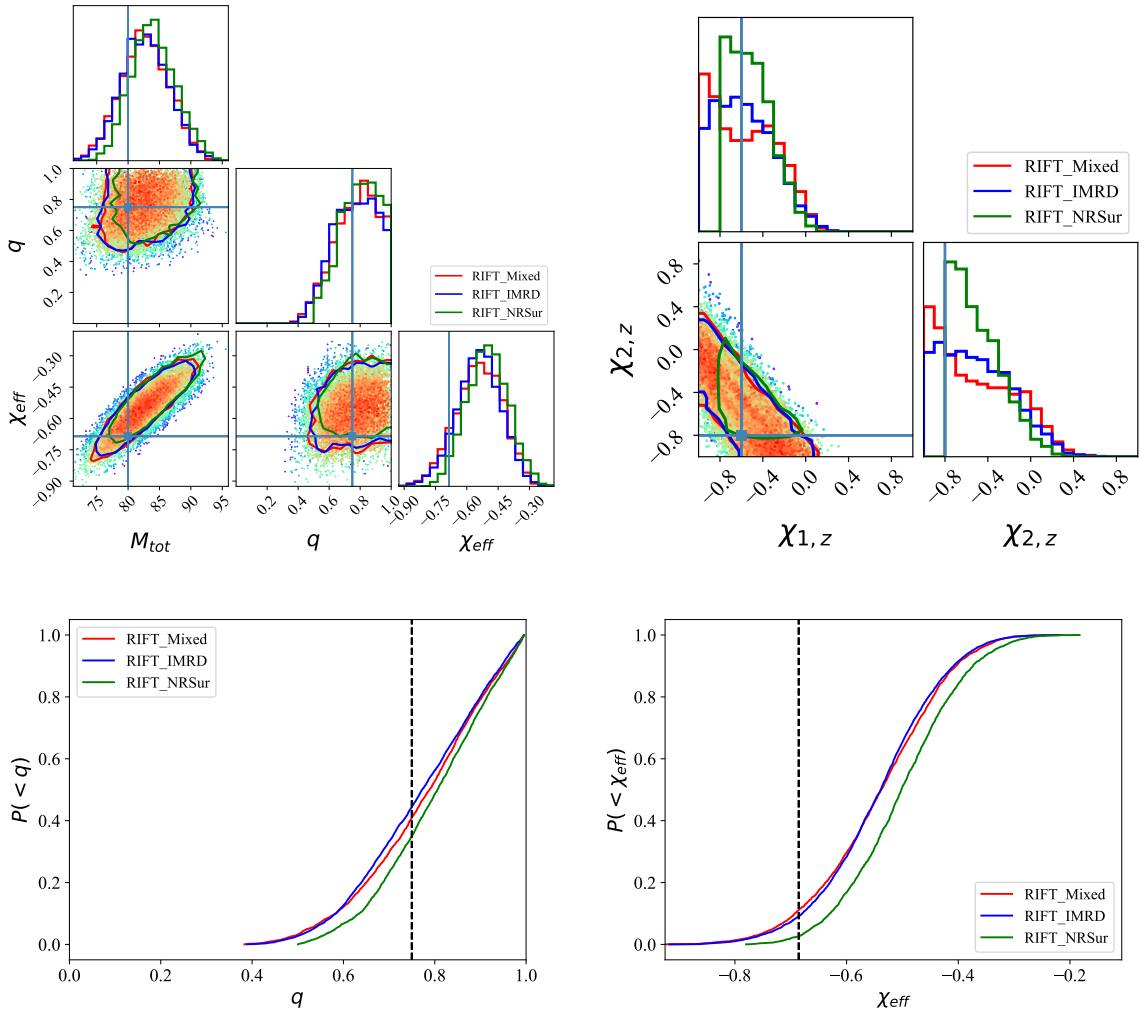

Figure 2.4: **Mixing models for a single analysis** This figure shows the inferred parameters for source RIT-2, in a format similar to Figure 2.1. The colored curves represent 90% credible intervals derived using IMRPhenomD (blue), NRSur7dq2 (green), and a mixture of models (red). In the mixture approach, we employ NRSur7dq2 with modes up to $\ell_{\rm max} = 2$ in its region of validity ($q > 0.5$ and $|\chi_{i,z}| < 0.8$) and IMRPhenomD elsewhere. Points indicate the locations of individual likelihood evaluations with the redder points representing a higher likelihood and bluer points representing a lower likelihood. The grid was originally produced from a run using SEOBNRv4.

($\ln \mathcal{L}_{\rm marg}$) enables detailed and easily-understood diagnostics as users change one feature of their calculation at a time (e.g., mode content; data conditioning or noise model).

Figure 2.5 shows a concrete illustration of this strategy, applied when interpreting the non-precessing binary source model RIT-2. In this case, parameter inference was performed using a fiducial model to generate a sequence of ever-finer evaluation grids. The net grid was then applied to two other models, leading to different predictions. Because these models all rely on the same input grid $\boldsymbol{\lambda}_\alpha$, we can directly diagnose which features drive differences in our posterior distributions. For example, in this case (Figure 2.5) this strategy helps us assess the relative role of restricted χ_i versus model differences such as higher modes in changing inferences about $\chi_{\rm eff}$ and hence M, q. In the previous example in particular (Figure 2.4), comparing the green (NRSur7dq2) and red (mixed) posterior distributions to the blue (IMRPhenomD) result, we see that the NRSur7dq2 posteriors are significantly impacted by the constraint that $|\chi_i| < 0.8$, a limit inherent in NRSur7dq2 . This limit strongly influences the $\chi_{\rm eff}$ distribution and, through correlations, the mass and mass ratio distribution. Once the domain of NRSur7dq2 is modestly extended to cover the whole parameter space, we find the mixed result (red) and IMRPhenomD result (blue) agree well with one another.

As another practical demonstration of this identical-grid strategy, we directly compare parameter inference using two nominally-equivalent models (SEOBNRv4 and SEOBNRv4_ROM) using two of the previously-presented synthetic sources: RIT-1 and RIT-2. Figure 2.6 shows our analysis of RIT-1, where the 1D $\chi_{\rm eff}$ distribution from these analyses; the dark red and red lines represent SEOBNRv4_ROM and SEOBNRv4 respectively. To demonstrate the visually apparent discrepancies are solely attributed to differences between the two underlying models, we repeated each analysis with LI; results are shown with dashed lines. Both RIFT and LI produce equivalent answers using the same model, but the posterior inferences produced using the two models differ. Figure 2.7 shows a similar analysis for RIT-2 two panels with the same 2D $\chi_{\rm eff}$ vs q distributions but with different colored $\ln \mathcal{L}$ grid. The dark red and red distributions represent the analyses using SEOBNRv4_ROM and SEOBNRv4 respectively. The left panel shows the $\ln \mathcal{L}$ grid when running with SEOBNRv4, and the right

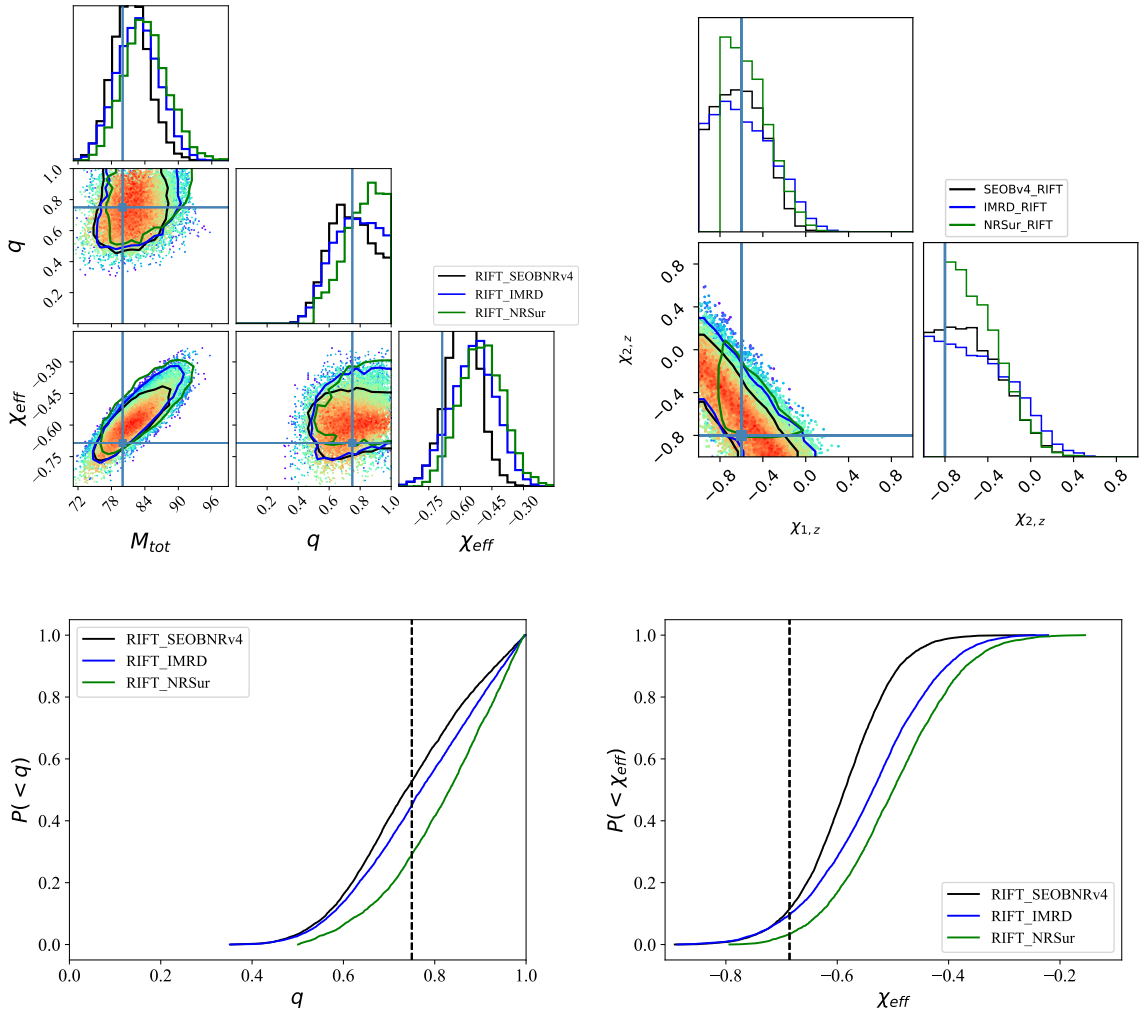

Figure 2.5: **Different models on the same parameter grid** This figure shows inferences about RIT-2, described in Table 2.1. The contours, points, and curves used in this figure are styled according to scheme described in Figure 2.1. Points indicate the locations of individual likelihood evaluations with the redder points representing a higher likelihood and bluer points representing a lower likelihood. The grid was originally produced from a run using SEOBNRv4 i.e. the black curve was produced by fitting these points.

m_1	m_2	χ_{1z}	χ_{2z}	$\chi_{\rm eff}$	$\Delta \ln \mathcal{L}$	$\ln \mathcal{L}_{\rm v4}$	$\ln \mathcal{L}_{\rm v4\_ROM}$	$Match$(v4, v4_ROM)
48.6	44.2	-0.01	-0.35	-0.17	5.98	154.2	160.2	0.997
47.3	44.1	-0.28	-0.15	-0.22	5.93	156.1	162.1	0.998
56.7	33.5	0.14	-0.95	-0.27	5.46	155.9	161.3	0.998
50.1	28.2	-0.80	-0.86	-0.82	5.45	158.2	163.6	0.998
50.3	39.1	-0.10	-0.59	-0.31	5.45	159.6	165.0	0.998
42.7	31.8	-0.99	-0.72	-0.88	5.40	162.1	167.5	0.998
48.9	39.7	-0.23	-0.27	-0.25	5.33	154.8	160.2	0.998
45.1	30.3	-0.89	-0.82	-0.86	5.29	161.4	166.7	0.998
58.0	37.4	-0.12	-0.30	-0.19	5.24	155.5	160.7	0.997
50.0	28.3	-0.85	-0.78	-0.82	5.11	158.6	163.7	0.998

Table 2.2: **Parameter grids with largest** $\Delta \ln \mathcal{L}$: This table shows the parameters, $\ln \mathcal{L}$, $\Delta \ln \mathcal{L}$, and the match between the two models at those points. The points are in order of decreasing $\Delta \ln \mathcal{L}$. Even though these points have a significant $\Delta \ln \mathcal{L}$, the match for all the points are relatively high.

panels shows the $\ln \mathcal{L}$ grid when running with SEOBNRv4_ROM. In both cases, the color scale indicates the value of $\ln \mathcal{L}$ where the redder points represent high $\ln \mathcal{L}$ and bluer represent lower $\ln \mathcal{L}$. By comparing these two panels, one can not only again see differences between the distributions but also differences between the underlying $\ln \mathcal{L}$ values produced with these two waveform approximations.

Because we used an identical grid with both approximations, we can examine the waveform parameters associated with the largest disagreements, illustrated in Table 2.2. This table provides the two $\ln \mathcal{L}$ values estimated with both waveforms, the difference $\Delta \ln \mathcal{L}$, and the match between the two models for precisely these waveform parameters. Despite the low amplitude, low mismatch, and short duration, these two analyses led to astrophysically noticeably different conclusions about important parameters (e.g., mass ratio and $\chi_{\rm eff}$) which, when biased, will impact population inferences. The detailed point-by-point analysis enabled by our method let us quickly demonstrate the impact of model systematic error in a regime where that systematic was previously presumed observationally irrelevant: a small systematic mismatch ($\simeq 3 \times 10^{-3}$), and sources and observing scenarios comparable to current observations. This example demonstrates the utility of our approach to identify sources of systematic error in GW parameter inference.

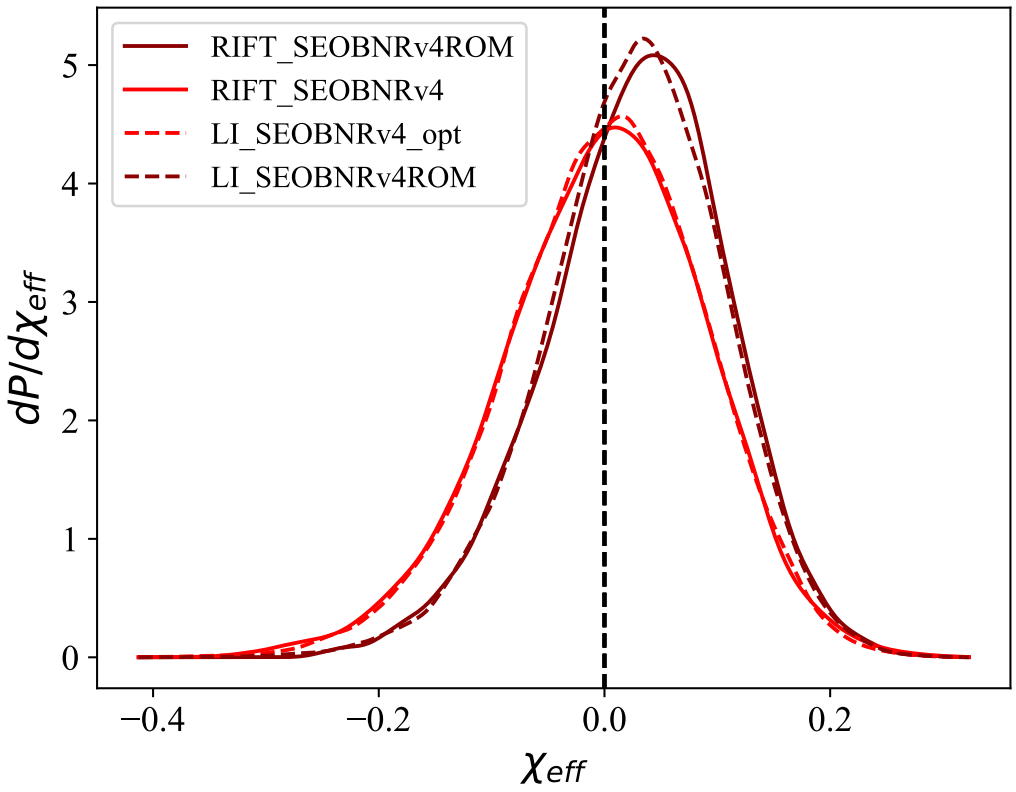

Figure 2.6: **Systematic Differences between SEOBNRv4 and SEOBNRv4_ROM - 1**: This figure shows the 1D $\chi_{\rm eff}$ distributions for analyses on RIT-1. The solid and dotted lines represent analyses done with RIFT and LI respectively. The dark red and red lines represent analyses run with SEOBNRv4_ROM and SEOBNRv4 respectively. This figure shows a minor but significant differences between these two models. By comparing the two different techniques with the two different models, it's clear that the differences are due to systematics between the two models and not differences in the PE techniques. See Table 2.2 for a point-by-point comparison of the 2 analyses.

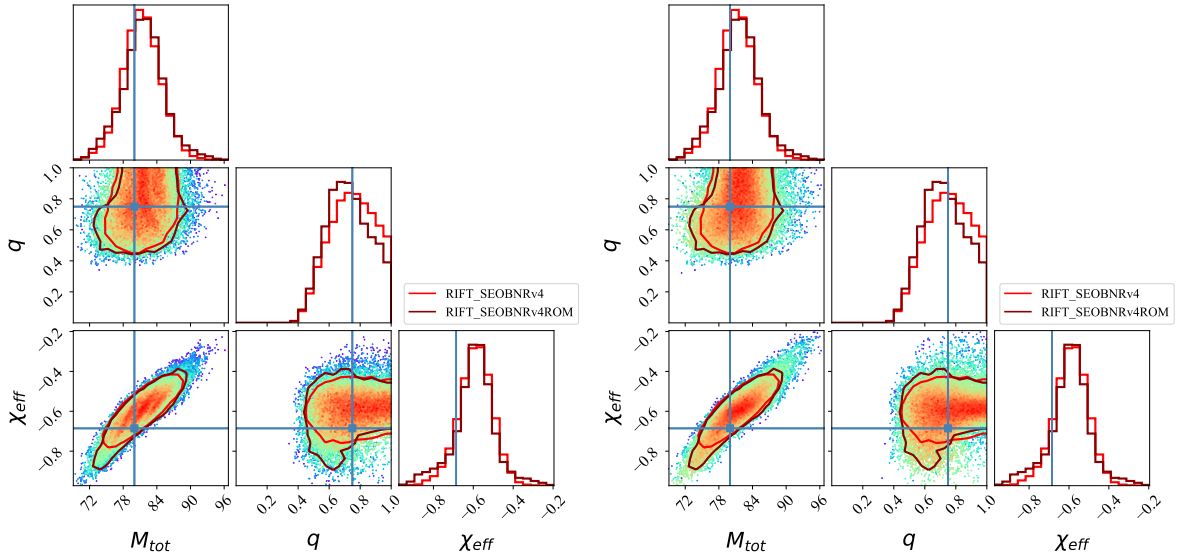

Figure 2.7: **Systematic Differences between SEOBNRv4 and SEOBNRv4_ROM - 2**: This figure shows the 2D $\chi_{\rm eff}$ vs q distributions for analyses on RIT-2. The dark red and red lines represent analyses run with SEOBNRv4_ROM and SEOBNRv4 respectively. The left panel shows the $\ln \mathcal{L}$ grid from the SEOBNRv4 analysis, and the right panel shows the $\ln \mathcal{L}$ grid from the SEOBNRv4_ROM analysis.The redder points represent a higher $\ln \mathcal{L}$ and the bluer points represent a lower $\ln \mathcal{L}$. By comparing these to grids, it is clear that the discrepancy between the two analyses is evident in the $\ln \mathcal{L}$ grids as well as in the distributions.

2.4.3 Astrophysical population inference

Since RIFT separates the process of estimating the likelihood and producing fair draws from the posterior, it is ideally suited for astrophysical inference, where all events are reassessed concurrently with different astrophysical priors. In this application, the ability to robustly explore alternative priors is paramount, particularly when future observations may motivate detailed reanalysis of earlier events. As described in Section 2.1 and Appendix 2.6, RIFT can efficiently generate posterior distributions using any prior whatsoever, by performing a suitable Monte Carlo.

In principle, other methods can also reanalyze existing results using alternative priors. Notably, with a sufficiently large list of posterior samples performed using a prior without compact support, the procedures described in Section 2.1 could produce weighted samples suitable to any prior. However, as emphatically demonstrated by all of our figures (Figs 2.2) and by Williamson et. al. [2], a finite list of samples from our posterior distribution with fiducial priors have compact support. For precessing binaries, previous analyses do not sample extremal spin. By contrast, by preserving the full (marginalized) likelihood, RIFT can evaluate the implications of a nearly arbitrarily extreme choice of prior, with minimal additional computational cost. In related work [111], we employ this approach with RIFT inputs for astrophysical inference.

2.4.4 Inferences about the nuclear equation of state

RIFT provides direct estimates of the (marginalized) likelihood $\ln \mathcal{L}_{\rm marg}(\boldsymbol{\lambda})$, allowing us to carry out subsequent calculations which require it. As a concrete example, we can more efficiently deduce the nuclear equation of state by post-processing generic calculations performed by RIFT without added computational expense; see, e.g., the analysis performed [112].

Specifically, in Section 2.3.3 and Figure 2.3 we carried out model-independent inference of binary neutron star masses, spins, and tidal deformabilities λ. Any proposed nuclear equation of state imposes a specific relationship between the (source-frame) neutron star gravitational masses m_i and the tidal λ parameter ($\lambda(m_i) = \lambda_i$). Imposing this relationship, we can re-

peat our parameter inference both to draw tighter inferences about binary parameters and to deduce how well the proposed equation of state matches the data, after marginalizing over all other quantities. With RIFT, this reanalysis requires a particularly efficient re-use of the marginalized likelihood $\ln \mathcal{L}_{\rm marg}(m_{1,z}, m_{2,z}, \lambda_1, \lambda_2, \chi_1, \chi_2)$ as a function of its parameters: the redshifted detector-frame masses $m_i(1+z) = m_{i,z}$; the two neutron star tidal deformabilities λ_i; and the two neutron star spins $\chi_i \in [-\chi_{\rm max}, \chi_{\rm max}]$, assumed parallel to the orbital angular momentum. Specifically, for each proposed equation of state, characterized by some hyper-parameters γ, we evaluate the a quantity equivalent to the evidence $\int d\boldsymbol{\lambda} \mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})$ appearing in Eq. (2.1.6):

$$I(\gamma) = \int dm_1 dm_2 d\chi_1 d\chi_2 p(\chi_1, \chi_2) dz p(z|m_1, m_2) \mathcal{L}_{\rm marg}(m_1(1+z), m_2(1+z), \lambda_1(m_1|\gamma), \lambda_2(m_2|\gamma), \chi_1, \chi_2) \quad (2.4.9)$$

where $p(z)$ is the posterior redshift distribution given a source with masses m_1, m_2 was observed, which can be efficiently extracted from RIFT or tightly constrained by electromagnetic observations. Appendix 2.8 describes techniques to efficiently evaluate this expression for generic observations. Just as with $\mathcal{L}_{\rm marg}$ and intrinsic binary parameters, we can evaluate the marginalized likelihood $I(\gamma)$ on a grid of equation of state parameters; interpolate; and generate posterior distributions over those equation of state parameters, for one or more events.

Figure 2.8 shows a concrete example of this procedure, applied to the spectral equation of state parameterization introduced by Lindblom [113, 114] and to a synthetic binary NS generated using the APR4 equation of state [115], as previously analyzed in Figure 2.3. For this low-redshift source, similar to GW170817, we will assume the redshift is well-determined, eliminating that factor in the integrand. [Equivalently, for this low-redshift source, the source-frame and detector-frame masses are sufficiently similar that $\lambda(m_{1,z}|\gamma) \simeq \lambda(m_1|\gamma)$ (i.e., the data can't discriminate between them) so the integral over redshift can be performed once and for all, and all remaining expressions carried out using redshifted masses; see Appendix 2.8.] As a concrete proof of concept, this figure shows the marginalized likelihood derived from this single event as a function of two spectral equation of state parameters, as well as Bayesian

inferences about those two equation of state parameters when holding others fixed. In this analysis, we include only causal equations of state ($v < c$) and require a maximum neutron star mass greater than $2M_\odot$, motivated by measurements of PSRs 1614-2230 and J0348+0432 [116, 117]. More generally, other observational and theoretical constraints will provide a prior $p(\gamma)$ on the equation of state parameters. In terms of this prior, the marginal posterior distribution on equation of state parameters is proportional to $I(\gamma)p(\gamma)$, and the marginal posterior distribution for any other expression of interest follows by either quadrature or use of weighted samples, as before. As a concrete example, if γ_α are n representative EOS configurations from the marginal posterior distribution and, for each α, we generate a Monte Carlo approximation to Eq. (2.4.9) and retain the weighted samples $(w_{k,\alpha}, \boldsymbol{\lambda}_{k,\alpha})$ for $k = 1 \dots N$ needed to evaluate it, then the marginal distribution of $\widetilde{\Lambda}$ [Eq. (1.1.24)] follows by Monte Carlo integration of the cumulative distribution: $P(< \widetilde{\Lambda}) \simeq (nN)^{-1} \sum_{k,\alpha} w_{k,\alpha} p(\gamma_k) \Theta\left(\widetilde{\Lambda} - \widetilde{\Lambda}(\boldsymbol{\lambda}_{k,\alpha})\right)$.

By providing marginalized likelihoods $I(\gamma)$, this approach to EOS inference enables the same powerful embarassingly-parallel and postprocessing-dominated approach used by RIFT itself. For example, we can combine inferences from multiple events by simply multiplying the associated likelihood distributions $[I(\gamma|d_1)I(\gamma|d_2)\dots]$. The end-user is free to efficiently adopt any prior of interest after the initial analysis; to assess whether choice of priors limits or dominates their analysis; and to incrementally extend their parameter space exploration with the minimum necessary computational expense. This approach is well-suited to rapidly-converging and low-dimensional parameterizations like the spectral method shown in Figure 2.8. However, due to the remarkably low cost of all integrals involved and our reliance on precomputed marginalized likelihoods, this approach can also be applied to assess generic EOS parameterizations inside standard Markov chain Monte Carlo algorithms. Though the above discussion is written as if equation of state dependence enters only through the tidal deformabilities λ, this procedure works when the underlying dynamics and radiation models include additional multipolar couplings. While similar parameterized equations of state have already been developed and used to infer the equation of state from individual binary NS GW measurements [118], our approach will make the best use of multiple observations, as we don't

need additional infrastructure or approximations to estimate $I(\gamma)$.

For low-redshift sources like GW170817 where $\lambda(m_z) \simeq \lambda(m)$ and hence where the redshift distribution $p(z|m_1, m_2)$ is not needed, and equivalently for sources with known host galaxy redshifts, we have also explored an alternative approach to compute $I(\gamma)$. In this approach, we evaluate $\ln \mathcal{L}_{\rm marg}$ on a grid of mass and spin choices, using the specified EOS to determine $\lambda_i = \lambda(m_i|\gamma)$. We then perform the usual RIFT approach to estimate $\ln \mathcal{L}_{\rm marg}$ and hence to compute $\int d\boldsymbol{\lambda} \mathcal{L}_{\rm marg}(\boldsymbol{\lambda}) p(\boldsymbol{\lambda})$, a byproduct of the Monte Carlo integration procedures used to generate our posterior distributions. While this approach can be useful for ranking the relativity validity of a few EOS, its duplicative computations and omission of redshift make it too burdensome for long-term use.

2.5 Analysis of a real event: GW150914

In this section, to describe one example of our method on real data, we analyze GW150914 with RIFT using several different approximate waveform models.[1] The analyses presented in this section were done with the C02 calibration frames with strain noise power spectral densities (PSDs) estimated from data segments near each event, similar to the PSDs used in [10]. As with all RIFT analyses, we assume zero calibration uncertainty. Gravitational wave strain data identical to the inputs employed here are available from the LIGO Open Science Center [119]. In these comparisons, we also employ different spin priors, such as the volumetric spin prior [Eq. (2.6.13)], than those employed by Abbott et al. [19, 120] in the original analysis of this event. Finally, to be consistent with previous analyses with numerical relativity simulations [9] and keeping in mind finite model duration, when using NRSur7dq2 we adopt a minimum frequency of 30Hz; however, we adopt a minimum frequency of 20Hz for our other analyses. The choice of lower cutoff frequency has little impact on our comparisons [9].

Figure 2.9 shows our analyses of GW150914 with RIFT using nonprecessing waveform models, all computed using a prior where that each $\chi_{i,z}$ is a uniform random number. As in

[1] Since our paper was submitted, RIFT was used in analysis of the GWTC-1 catalog.

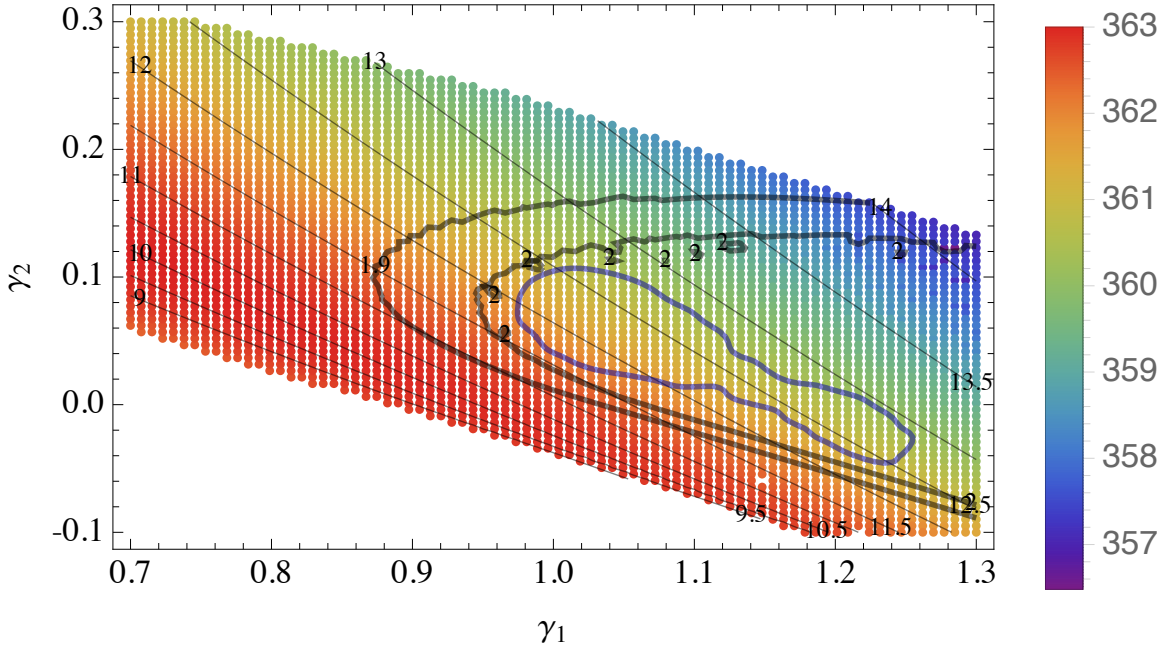

Figure 2.8: **Inferences about the equation of state**: Colors indicate the marginalized likelihood $I(\gamma)$ versus the two parameters γ_1, γ_2 of the Lindblom et al spectral EOS representation, applied to a binary neutron star source similar to Tidal-1. For context, the thin gray lines show contours of constant radius of a 1.5 $M_\odot$ neutron star, according to this equation of state $[R(m = 1.5M_\odot)]$; the two heavy black lines show contours of constant maximum mass for $2M)_\odot$ and $1.9M_\odot$, respectively. For the purposes of this proof-of-concept calculation, we explore only these two variables, fixing the remaining spectral EOS parameters to $\gamma_3 = \gamma_4 = 0$, $p_o = 2.272 \times 10^{33}\mathrm{dyne/cm}^2$, $x_{\max} = 7.25$, and $\epsilon_o/c^2 = 2.05 \times 10^{14}\mathrm{g/cm}^3$. The solid blue line shows the 90% credible interval on the inferred equation of state, after restricting to a causal EOS and restricting the maximum mass to be greater than $1.97M_\odot$.

previous result derived by directly comparing GW150914 to numerical solutions of Einstein's equations [9], we find that our posterior inferences derived a waveform model that includes higher modes (here, NRSur7dq2 with $\ell \leq 3$) more sharply constrains some parameters like the binary mass ratio q. However, our analysis and previous results [9] suggest that, in an aligned-spin analysis of this event which adopts a prior not strongly disfavoring anti-aligned spins, the posterior has substantial support outside of the domain of validity of NRSur7dq2. As demonstrated previously in Figure 2.4, the domain of validity of NRSur7dq2 can impact inferences if the posterior support extends to the edge of the model domain. In a companion study devoted to direct comparison to numerical relativity solutions, we will more carefully address differences between this result and the analysis presented in [9], using numerical relativity simulations to fill the gaps.

Next, to provide a comparison that allows for all spin degrees of freedom, we analyze GW150914 with SEOBNRv3 [121], NRSur7dq2 ($\ell \leq 3$), and IMRPhenomPv2. Figure 2.10 shows the 90% credible intervals for the RIFT_SEOBNRv3, RIFT_NRSur7dq2 ($\ell \leq 3$), and RIFT_IMRPhenomPv2 analyses in black, green, and blue, respectively. Note that the SEOBNRv3 analysis was done with different settings; therefore, some differences are expected. First and foremost, consistent with our and other prior work, we see differences between the posterior inferences derived using different waveform approximations and assumptions, even for sources like GW150914 with now-fiducial binary parameters and source amplitudes. Second, consistent with an earlier investigation of GW150914 which directly compared it to numerical relativity simulations [9], we see that including higher harmonics enables us to draw sharper conclusions about binary parameters – here, the binary mass ratio. Our analysis differs from that prior investigation work in two key ways: precessing inference and χ limits. While the previous study [9] compared GW150914 to generic simulations, a posterior distribution was estimated only on the basis of *nonprecessing* simulations; by contrast, our analysis employs a fully precessing model. Conversely, our use of NRSur7dq2 is limited to $|\chi_i| < 0.8$, which directly constraints our ability to draw generic inferences about BH spins.

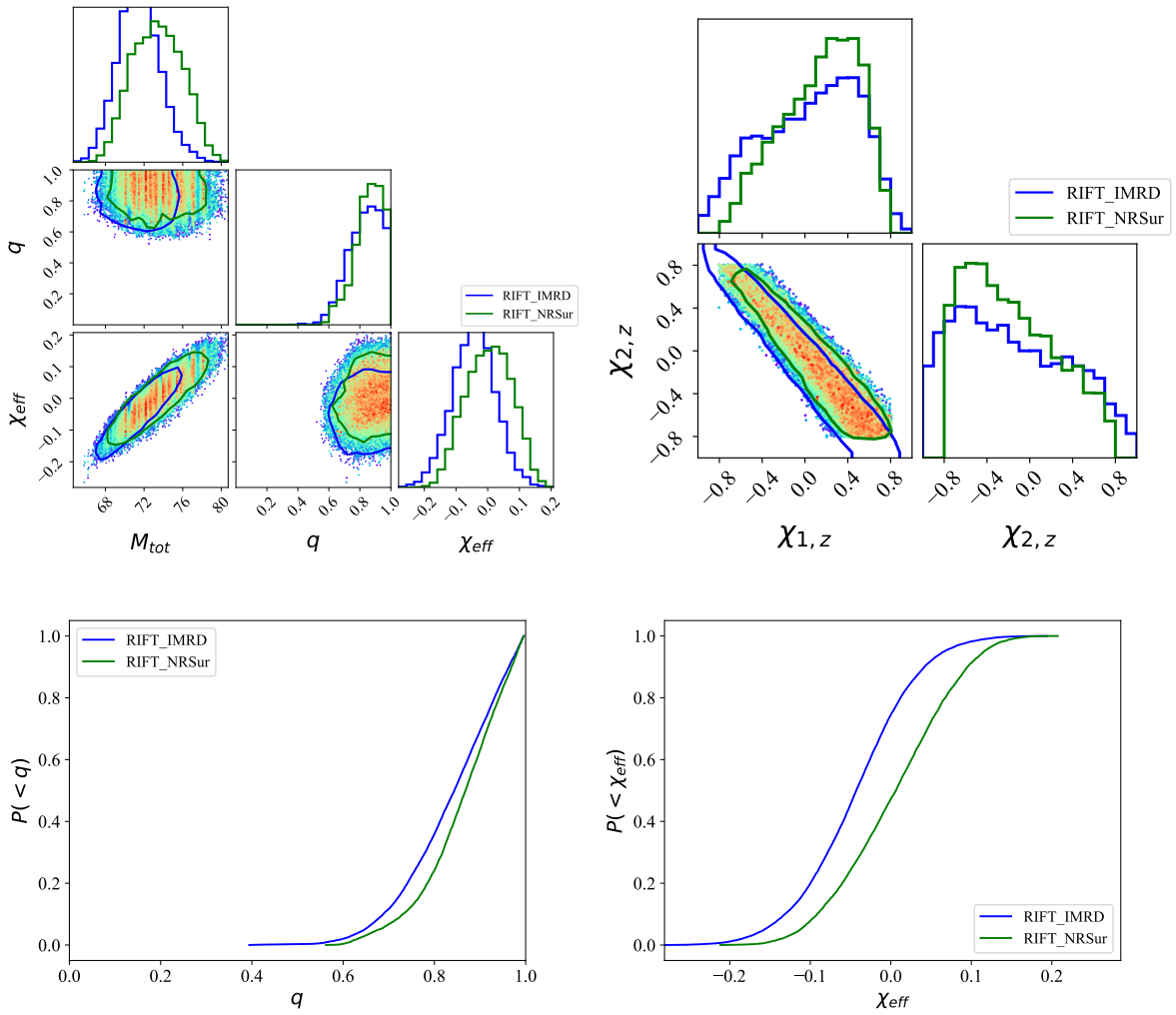

Figure 2.9: **Reanalysis of GW150914: Aligned models** - This figure shows the results of a reanalysis of the first GW detection GW150914, using a prior where each $\chi_{i,z}$ is a uniform random number and the assumption that both spins must be parallel to the orbital angular momentum. The top two panels show one- and two-dimensional marginal distributions in redshifted total mass $M_{\rm tot}, q, \chi_{\rm eff}$ and in the components of each BH's spin $(\chi_{i,z})$ along the orbital angular momentum. The bottom panel provides one-dimensional marginal cumulative distributions for q and $\chi_{\rm eff}$, to highlight differences between different waveforms. Curves are colored and styled following the convention adopted in previous figures; in particular, the solid green lines to NRSur7dq2 with the spins restricted along the orbital angular momentum with all modes up to $\ell_{\rm max} = 3$, and blue lines to IMRPhenomD. The top right and hence bottom left panels show significant differences between NRSur7dq2 and other approximations principally due to the range of validity of the NRSur7dq2 model, which is valid only for $q > 0.5$ and $|\chi_i| < 0.8$. Note that the RIFT_SEOBNRv3 run was run with a different configuration than the other two analyses.

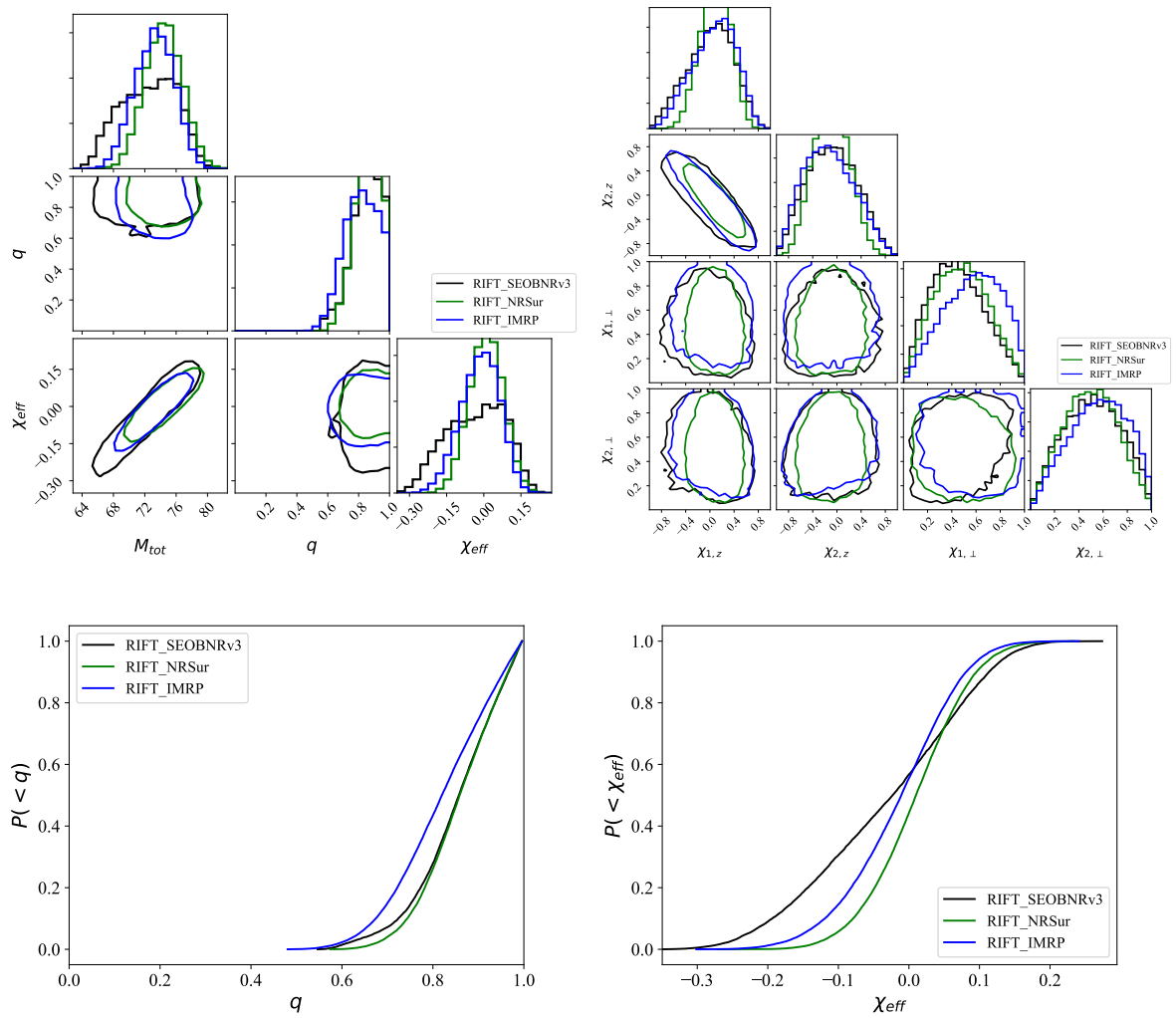

Figure 2.10: **Reanalysis of GW150914: Precessing models** - This figure shows the results of a reanalysis of the first GW detection GW150914, using the volumetric spin prior [Eq. (2.6.13)] and waveform models that allow for spin-orbit misalignment. The top two panels show one- and two-dimensional marginal distributions in redshifted total mass $M_{\rm tot}, q, \chi_{\rm eff}$ and in the components of each BH's spin along ($\chi_{i,z}$) and perpendicular ($\chi_{i,\perp}$) to the orbital angular momentum. The bottom panel provides one-dimensional marginal cumulative distributions for q and $\chi_{\rm eff}$, to highlight differences between different approaches. Curves are colored and styled following the convention adopted in previous figures; in particular, black lines correspond to SEOBNRv3, solid green lines to NRSur7dq2 with fully precessing spins with all modes up to $\ell_{\rm max} = 3$, and blue lines to IMRPhenomPv2. The top right panel shows significant differences between NRSur7dq2 and other approximations principally due to the range of validity of the NRSur7dq2 model, which is valid only for $q > 0.5$ and $|\chi_i| < 0.8$. Note that the SEOBNRv3 analysis was done with different settings; therefore, some differences are expected.

2.6 Appendix A: Coordinate systems and priors

In this appendix, we summarize the coordinate systems and priors adopted to perform the fit and in particular carry out the Monte Carlo integration procedure for $\ln \mathcal{L}_{\rm marg}$ which produces weighted posterior samples, as described in Section 2.1. The design structure of the adaptive Monte Carlo integrator developed for ILE and re-used in our code drives several prior and coordinate choices. At a very low level, the adaptive procedure assumes the initial true prior p and (adapted) sampling prior p_s are both separable: $p(x_1 \dots x_n) = p(x_1)p(x_2)\dots p(x_n)$ for $\{x_1 \dots x_n\}$ the coordinate system used to perform the Monte Carlo. Moreover, because the adaptive procedure optimizes each sampling prior p_s in one coordinate at a time, the overall algorithm's performance dramatically improves when the posterior is compatible with this coordinate system (i.e., the posterior approximately fills a sub-hypercube of our computational domain). Whenever we have confident reason to suspect a particular parameter does not impact the posterior distribution whatsoever, we can analytically marginalize over superfluous parameters and perform a fit which omits these parameters. Finally, we can explore alternative priors by using weighted posterior samples, using the ratio of new prior relative to reference prior in all coordinate dimensions.

2.6.1 Masses

We adopt a uniform prior over the (detector-frame) masses m_1, m_2 (with $m_1 \geq m_2$), which can be expressed in several equivalent coordinate systems by a suitable Jacobian transformation. Following past practice, we nominally employ a triangular region with $m_i \geq m_{min}$ and $M = m_1 + m_2 \leq M_{tot,max}$; the prior is $p(m_1, m_2) = 2/(M_{\rm max} - m_{min})^2$, if both regions $m_1 > m_2$ and $m_1 < m_2$ are included in the integration, and twice that otherwise. In most cases, we perform our underlying calculations in $\mathcal{M}, \eta$ coordinates, where $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ and $\eta = m_1 m_2/(m_1 + m_2)^2$. In these coordinates, the prior can be represented as [22]

$$p(\mathcal{M}, \eta) = \frac{4}{(M_{\rm max} - m_{min})^2} \frac{\mathcal{M}}{\eta^{6/5}\sqrt{1 - 4\eta}} \,. \tag{2.6.10}$$

As this prior diverges at $\eta = 1/4$, it is not well-suited to posterior distributions with significant support very close to the equal mass line, like binary neutron stars. For such cases, we change coordinates to $\delta = \sqrt{1-4\eta}$; the prior becomes

$$p(\mathcal{M}, \delta) = \frac{16 \times 2^{2/5}}{(M_{\rm max} - m_{min})^2} \frac{\mathcal{M}}{(1-\delta^2)^{6/5}} \ . \tag{2.6.11}$$

In rare cases, we may employ coordinates M and $q = m_2/m_1$; in this case, prior is [95]

$$p(M, q) = \frac{4}{(M_{\rm max} - m_{min})^2} \frac{M}{(1+q)^2} \ . \tag{2.6.12}$$

For confident detections, the mass posterior will not have significant support outside a compact region. Because we can freely change the prior in regions far outside the posterior and not impact results, we often adopt technically convenient boundaries. For example, rather than transform the triangular region of permissible masses m_1, m_2 to a more complicated boundary in $\mathcal{M}, \eta$, we simply adopt suitable intervals in $\mathcal{M}$ and η which contain the posterior.

2.6.2 Spins

The spin degrees of freedom are characterized by the spin angular momenta $\mathbf{S}_i$ or equivalently their dimensionless spins χ_i/m_i^2. For most astrophysically relevant scenarios, the posterior distribution strongly constrains $\chi_{\rm eff}$ [Eq. (1.1.22)] but not the complementary combination of aligned spins. Lacking a compelling astrophysically-motivated choice, several spin-dependent priors have been adopted, which employ different numbers of spin degrees of freedom.

Volumetric spin prior (generic, cartesian): By default, we employ a prior which is uniform in each component of the dimensionless spins χ_i in some hypercube: $\chi_i \in [-\chi_{\rm max}, \chi_{\rm max}]$. After generating posterior samples, we then eliminate samples with either dimensionless spin above our threshold ($|\chi| > \chi_{\rm max}$). Combined, our hypercube-plus-cut procedure produces an effective prior that is uniform over the volume of a coordinate sphere of radius $\chi_{\rm max}$:

$$p_{vol}(\chi_i, \theta_i, \phi_i) = \frac{3}{4\pi\chi_{\rm max}^3} \chi_i^2 \sin\theta_i \tag{2.6.13}$$

where χ_i, θ_i, ϕ_i are the spin vector's polar coordinates relative to the orbital angular momentum vector [122].

Beta function spin prior (precessing): More generally, we can adopt a Beta-distribution prior on the spin magnitude

$$p_\beta(\chi_i, \theta_i, \phi_i | p, q) = \frac{\Gamma(2+p+q)}{\Gamma(1+a)\Gamma(1+b)} \frac{\chi_i^p (\chi_{\rm max} - \chi_i)^q}{4\pi \chi_{\rm max}^{p+q+1}} \sin\theta_i \tag{2.6.14}$$

In this representation, the volumetric spin prior corresponds to $p = 2, q = 0$. The default prior adopted by LIGO when inferring BH spins uses a uniform spin magnitude distribution [23], corresponding to $p = q = 0$:

$$p_{umag}(\chi_i, \theta_i, \phi_i) = \frac{1}{4\pi \chi_{\rm max}} \sin\theta_i \tag{2.6.15}$$

Aligned spin: When both spins must be parallel to the orbital angular momentum, by default we employ a uniform prior on $\chi_{i,z} \in [-\chi_{\rm max}, \chi_{\rm max}]$. To enable comparison to precessing LIGO results, we also employ an alternative prior ("zprior"), proposed by Veitch [123]

$$p_{zprior}(\chi_{i,z}) = \frac{1}{2\chi_{\rm max}} (-\ln|\chi_{i,z}/\chi_{\rm max}|) \tag{2.6.16}$$

which is equivalent to the uniform spin magnitude prior after marginalizing out other degrees of freedom. To demonstrate that p_{zprior} is the corresponding marginal distribution of Eq. (2.6.15), we perform the marginal integral, for convenience denoting spin components for convenience by x, y, z and $\chi_{\rm max}$ by R. The cumulative distribution $P(<z)$ can be evaluated by the following expression when $z < 0$:

$$P(<z) = \int_{<z} \frac{dr}{R} \frac{d\phi d\cos\theta}{4\pi} \tag{2.6.17}$$

$$= \frac{1}{2R} \int_{-1}^{|z|/R} d\cos\theta \int_{|z|/|\cos\theta|}^{R} d\cos\theta \tag{2.6.18}$$

$$= \frac{1}{2R}[z + R - z\ln|z/R|] \tag{2.6.19}$$

and the result follows by differentiation.

Correlated, separable spin priors: In almost all cases, the posterior distribution tightly constrains $\chi_{\rm eff}$ [Eq. (1.1.22)] but leaves the complementary degree of freedom almost completely unconstrained. To accelerate sampling in the common case with $m_1 \simeq m_2$, we can adopt a uniform spin prior in $\chi_{z,\pm} = (\chi_{1,z} \pm \chi_{2,z})/2$, then eliminate samples which have $\chi_{z,i}$ otherwise inconsistent with the limit imposed by $\chi_{\rm max}$. This approach can be directly employed to accelerate sampling with both the aligned uniform spin prior and the precessing volumetric prior; results for non-separable priors follow by reweighting posterior samples.

Marginal prior for $\chi_{\rm eff}$ (uniform): Because the individual spin components $\chi_{i,z}$ are rarely observationally accessible, one can imagine marginalizing out the superfluous degree of freedom, reducing the marginal likelihood and related Monte Carlo calculations to integrals like $I = \int_{|\chi_1|,|\chi_2|<1} f(\chi_{\rm eff}) d\chi_{1,z} d\chi_{2,z}$ via

$$I = \int f(\chi_{\rm eff}) d\chi_1 d\chi_2 = 4 \int f(\chi_{\rm eff}) p(\chi_{\rm eff}|q) d\chi_{\rm eff} \tag{2.6.20}$$

where for simplicity we adopt a uniform spin magnitude prior. This expression is not directly applicable to our low-level Monte Carlo technique, as this spin prior depends on mass ratio. Such an expression will however be a useful reference when we want to rescale posterior samples to alternative prior distributions.

Reweighting is only successful if the posterior has broad support. For nonprecessing inference, while analyses performed with Eq. (2.6.16) generally have support concentrated near to zero spin, an analysis with the uniform spin magnitude prior will have support generally for all $\chi_{i,z}$ consistent with the likelihood. When constructing a fiducial marginal $\chi_{\rm eff}$ distribution, we therefore derive it under the assumption of uniformly distributed $\chi_{i,z} \in [-\chi_{\rm max}, \chi_{\rm max}]$. In the interests of clarity and without loss of generality – all results scale linearly with $\chi_{\rm max}$ – in the derivation that follows we adopt $\chi_{\rm max} = 1$.

To evaluate $p(\chi_{\rm eff}|q)$, we first define a helpful shorthand (to avoid ambiguity)

$$g(a,b) = \frac{m_1 a + m_2 b}{m_1 + m_2} = (a + qb)/(1+q) \tag{2.6.21}$$

which is $\chi_{\rm eff}$. The integrand and prior has four natural breakpoints at $g(\pm 1, \pm 1)$, ordered so $g(-1,-1) \leq g(-1,1) \leq g(1,-1) \leq g(1,1)$. Within each region, we can do the integral $\int d\chi_1 d\chi_2 \delta(z - g(\chi_1, \chi_2))$ simply by keeping track of the limits of integration:

$$J(z) = \int_{-1}^{1} d\chi_1 \int_{-1}^{1} d\chi_2 \delta(z - g(\chi_1, \chi_2)) \tag{2.6.22}$$

$$= \int_{\chi_2 = \chi_2(\chi_1, \chi_{\rm eff})} d\chi_1 (1+q)/q \tag{2.6.23}$$

$$= \frac{(1+q)}{q} [\chi_{1,+} - \chi_{1,-}] \tag{2.6.24}$$

where $\chi_{1,\pm}$ are the largest and smallest allowed values of χ_1 for a given choice of $\chi_{\rm eff}$. Looking at the square, when $\chi_{\rm eff} < g(-1,1)$, we know $\chi_{1,-} = -1$ and when $\chi_{\rm eff} > g(1,-1)$ we know the upper bound is 1. Otherwise, we know $\chi_{1,\pm}$ occurs when $\chi_2 = \pm 1$, implying

$$\chi_{1,\pm}(z) = \pm q + z(1+q) \tag{2.6.25}$$

We find the following expression for our marginal prior

$$p(z|q) = \frac{1+q}{4q} \times \begin{cases} 1 - \chi_{1,-} & z \in [g(1,-1), g(1,1)] \\ \chi_{1,+} - \chi_{1,-} = 2q & z \in [g(-1,1), g(1,-1)] \\ \chi_{1,+} + 1 & z \in [g(-1,-1), g(-1,1)] \end{cases} \tag{2.6.26}$$

2.6.3 Tides

The tidal deformability of each compact binary can be characterized by a dimensionless parameter Λ_i, which is zero for BHs. By default, we adopt a uniform prior on $\Lambda_i \in [0, \Lambda_{\rm max}]$.

The leading-order effects of tidal deformation enter into the gravitational wave signal via two quantities $\widetilde{\Lambda}, \delta\widetilde{\Lambda}$ [Eq. 1.1.24] In most cases of current astrophysical interest, $\widetilde{\Lambda}$ can be weakly constrained and $\delta\widetilde{\Lambda}$ cannot be constrained at all.

Because of the constraint that $\Lambda_i \geq 0$, a corner in tidal parameter space, the marginal distribution of $\widetilde{\Lambda}, \delta\widetilde{\Lambda}$ near $\widetilde{\Lambda} \simeq 0$ increases linearly with $\widetilde{\Lambda}$. To more transparently reflect

the astrophysical significance of the posterior distribution of $\widetilde{\Lambda}$, it is helpful to adopt a prior corresponding to a *uniform* distribution of $\widetilde{\Lambda}$. Once again, such an alternative prior depends on binary mass ratios and therefore is useful only for post-processing and reweighting, not as part of our initial Monte Carlo analysis which requires separable priors. The underlying calculation of the marginal prior on $\widetilde{\Lambda}$ follows exactly like the calculation for $\chi_{\rm eff}$ above.

2.6.4 Prior recovery

When $\mathcal{L}$ is constant, our likelihood fitting and posterior generation code (CIP) must recover the desired input priors. As described in the text, these priors can be applied in two ways: at the lowest level, when performing the Monte Carlo integral itself; and by reweighting. Reweighting is frequently applied to achieve the spin priors preferred by the LVC (i.e., uniform in spin magnitude). Figure 2.11 demonstrates that CIP recovers the desired spin prior when operating in configurations needed to reproduce fiducial LVC results. In this figure, the settings adopted for CIP were chosen to produce Monte Carlo errors only marginally larger than the linewidths used in our figure, for presentation purposes.

2.7 Appendix B: Supplementary validation studies of iterative posterior generation

In this section we describe systematic tests of the general algorithm and specific implementation described in Section 2.1 to reconstruct the posterior distribution by means of iteratively fitting the likelihood distribution, then drawing candidate points from the posterior distribution.

In the first set of controlled tests, we generate synthetic likelihood function on a hypercube $x \in [-1,1]^d$ and a gaussian mixture model:

$$\mathcal{L}(x) = \mathcal{L}_{\rm ref} p(x) = \mathcal{L}_{\rm ref} \sum_k \frac{w_k}{(2\pi\sigma^2)^{d/2}} e^{-(x-\mu_k)^2/2\sigma^2} \tag{2.7.27}$$

where w_k are weights with $\sum_k w_k = 1$. We drew random weights; random gaussian centers

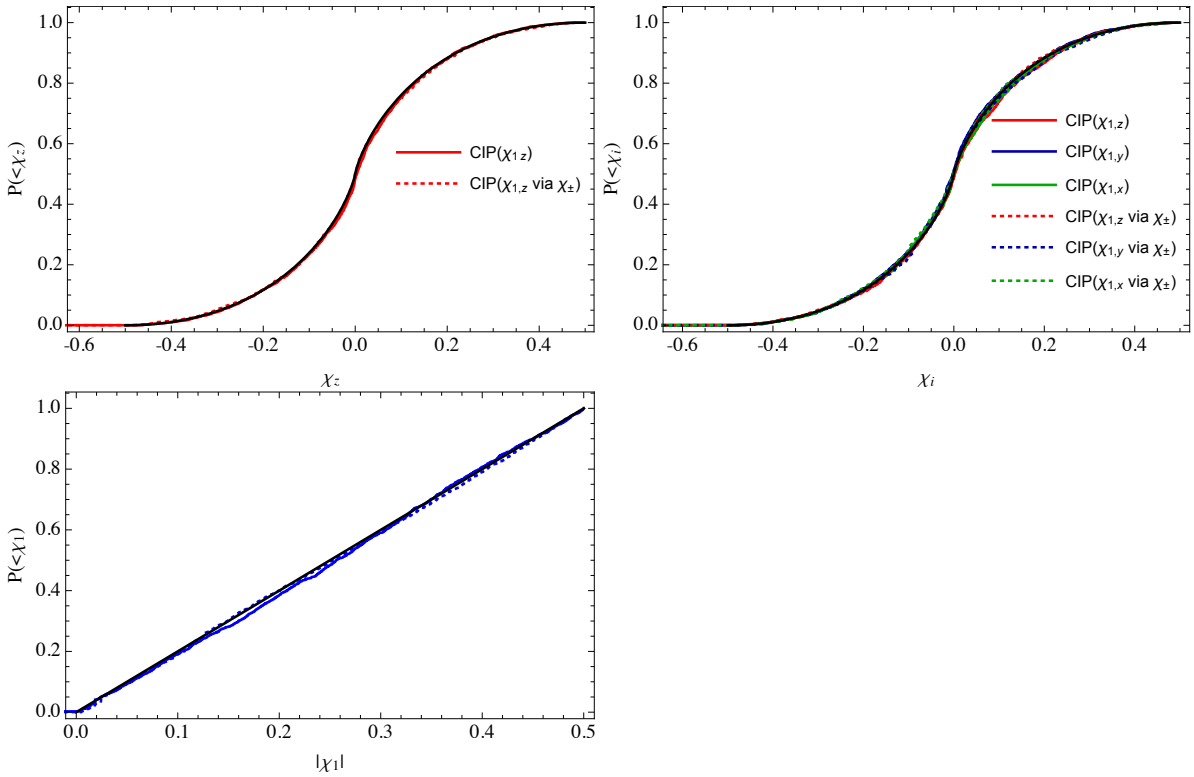

Figure 2.11: **Prior recovery**: *Top-left panel*: Comparison between analytic CDF predictions (black, solid) for $P(<\chi_{1,z})$ based on Eq. (2.6.16) and the corresponding empirical cumulative distribution of samples produced by CIP in two different circumstances with $\mathcal{L} = 1$ that should produce identical results. In this example, we adopt $\chi_{\rm max} = 0.5$. In this nonprecessing example, we request that CIP draw each $\chi_{i,z}$ directly from this prior (red). Alternatively, we can produce the same distribution (red, dotted) by jointly drawing both $\chi_{z,\pm}$ from a jointly uniform distribution, truncating to the desired range, then reweighting to obtain the desired one-dimensional marginal distribution. *Top-right panel*: In this precessing example, we request that CIP draw each $\boldsymbol{\chi}_i$ uniformly over a unit cube of edge length $2\chi_{\rm max}$; require that $|\boldsymbol{\chi}_i| < \chi_{\rm max}$; then rescale weighted Monte Carlo samples to obtain a uniform magnitude spin magnitude prior. The red, blue, and green curves show the corresponding cumulative distributions for $\chi_{1,z}, \chi_{1,y}, \chi_{1,x}$; all agree with the expected result. The corresponding dotted curves show the results of a similar calculation, except that the aligned-spin components $\chi_{i,z}$ were generated by drawing $\chi_{z,\pm}$ from a jointly uniform distribution, solving for $\chi_{i,z}$, and reweighting. *Bottom-left panel*: The distribution of $|\chi_1|$, calculated using the two procedures adopted for the center panel. The solid black line shows the expected distribution $P(<\chi_1) = \chi_1/\chi_{\rm max}$, while the solid and dotted blue lines show the results of the two procedures described above.

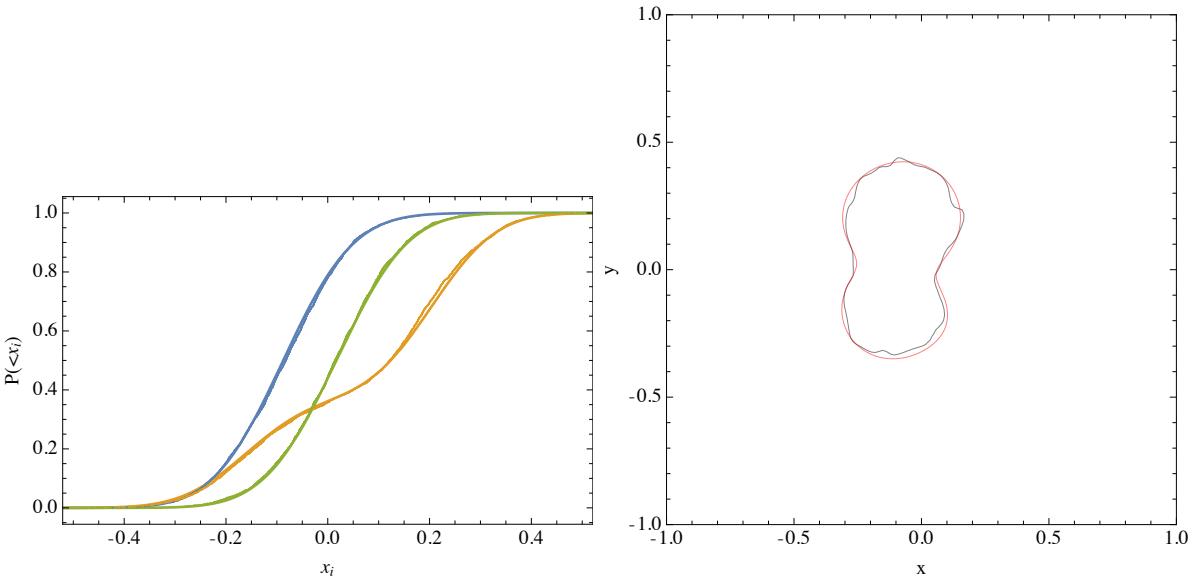

Figure 2.12: **Reliable reconstruction in three dimensions**: Parameter inference of a 3 dimensional gaussian mixture model with 4 components, compared with the analytic posterior distribution. *Left panel*: The *marginal, 1-dimensional* posterior estimates extracted from a mixture of four randomly-centered 3-dimensional gaussians. Solid lines show the cumulative 1d distributions for each of the three dimensions; points show the estimated distributions from our code.The numbers on the left provide an (approximate) KL divergence between each marginal 1d distribution and the known 1d marginal distribution. The first number refers to the first (blue) curve, and so forth (yellow, green, ...). *Right panel*: The 90% credible interval derived from the exact (red) and approximate (black) two-dimensional marginal distribution. The black curve is estimated based on 2000 samples drawn from the posterior distribution.

μ_k in the hypercube $[-0.6, 0.7]/\sqrt{d}$; and for convenience fixed $\sigma = 0.1$. We drew initial points randomly from the hypercube, seeded by a few points from the true posterior distribution, then applied our iterative code, iterating five times. Figure 2.12 shows an example of the output of our code, compared to the analytic one- and two-dimensional posterior distributions for a generic three-dimensional and four-component gaussian mixture model. A fairly accurate reconstruction will have $D_{KL} \lesssim 2\times10^{-2}$. For arbitrary gaussian mixture models in dimensions $d \leq 6$, we confirmed our approach consistently reproduces one- and two-dimensional posterior marginal distributions, with small error.

In the second test, we generate three sets of probability-probability plots (usually denoted PP plot), to assess the end-to-end performance of all aspects our method. Specifically, we generate many synthetic signals for 3-detector networks, with masses drawn uniformly in m_i

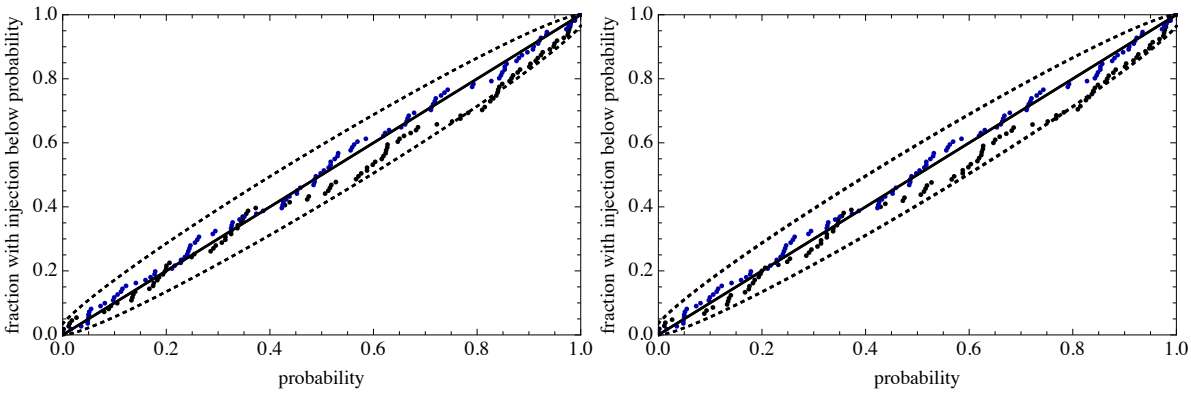

Figure 2.13: **Probability-probability plot**: For a set of synthetic BBH signals analyzed end-to-end with RIFT, a plot of $\hat{P}(\lambda_k)$ versus k/n for $k = 0 \ldots 1 - 1/n$, for the parameters $\mathcal{M}$ (black) and q (blue). The dotted lines show the standard 2-standard-deviation credible intervals implied by the binomial distribution for $\hat{P}$. *Left panel*: Analysis using SEOBNRv4 for synthetic sources and recovery. *Right panel*: Analysis using SEOBNRv4HM for synthetic sources and recovery.

in the region bounded by $\mathcal{M}/M_\odot \in [30, 60]$ and $\eta \in [0.2, 1/4]$ and with extrinsic parameters drawn uniformly from the priors described earlier, except that source distances are drawn between 1.5Gpc and 5Gpc. For simplicity, all of our synthetic sources have zero spin. (For complete reproducibility, we use `SEOBNRv4` and `SEOBNRv4HM`, starting the signal evolution at 8Hz but the likelihood integration at 20Hz, performing all analysis with 4096Hz timeseries in Gaussian noise with known PSDs. For the first pair of tests, for each synthetic event and for each interferometer, the same noise realization is used for both waveform approximations. Differences between the analyses therefore arise solely due to waveform systematics.) Using RIFT on each source k, with true parameters λ_k, we estimate the fraction of the posterior distributions which is below the true source value $\lambda_{k,\alpha}$ $[\hat{P}_{k,\alpha}(< \lambda_{k,\alpha})]$ for each intrinsic parameter α, again assuming all sources have zero spin. After reindexing the sources so $\hat{P}_{k,\alpha}(\lambda_{k,\alpha})$ increases with k for some fixed α, Figure 2.13 shows a plot of k/N versus $\hat{P}_k(\lambda_{k,\alpha})$ for both mass parameters. The recovered probability distribution is consistent with $P(< p) = p$, as expected.

2.8 Appendix C: Evaluating the evidence for an equation of state

In Section 2.4.4 we described a simple integral over source intrinsic parameters and redshift [Eq. (2.4.9)] to assess the compatibility with a given observation and a proposed equation of state, characterized by hyper-parameters γ which determine a relationship between tidal deformability λ and source-frame NS gravitational mass m: $\lambda = \lambda(m|\gamma)$. Because GW measurements naturally very tightly constrain the redshifted chirp mass $\mathcal{M}_z = \mathcal{M}(1+z)$ to a characteristic range of scale $\sigma_{\mathcal{M}_z}$, the integrand is nearly zero except for a very narrow range of $\mathcal{M}_z$ centered on $\mathcal{M}_{z,*}$. Additionally, to an excellent approximation the integrand does not depend on $\delta\widetilde{\Lambda}$, and the function $\lambda(m)$ is effectively constant when the mass changes by of order $\sigma_{\mathcal{M}_z}$. We therefore change variables to $\mathcal{M}_z, \eta, z, \chi_i$ [Eq. (2.6.10)], then perform the integral over $\mathcal{M}_z$ and χ_i:

$$I(\gamma) = \int dz d\eta \frac{4p(z|\mathcal{M}_{z,*})G\left(\eta, \widetilde{\Lambda}(\eta, \mathcal{M}_{z,*})\right)}{(1+z)^2(M_{\rm max} - m_{min})^2\eta^{6/5}\sqrt{1-4\eta}} \tag{2.8.28}$$

$$G(\eta, \widetilde{\Lambda}) = \int d\mathcal{M}\mathcal{M}\mathcal{L}(\mathcal{M}, \eta, \widetilde{\Lambda}) d\chi_1 d\chi_2 p(\chi_1, \chi_2) \tag{2.8.29}$$

where $\widetilde{\Lambda}_*(\eta, z)$ follows by evaluating $\lambda_i(m_i)$ using the masses m_i derived from η and the appropriate source-frame chirp mass $\mathcal{M}_{z,*}/(1+z)$, and where $p(z|\mathcal{M}_{z,*})$ is the fully marginalized distance distribution implied by this observation for a source with detector-frame chirp mass $\mathcal{M}_{z,*}$. Following Eq. (2.6.11), changing coordinates to δ rather than η removes the integrable singularity at $\eta = 1/4$ and makes numerical methods more robust.

Chapter 3

Testing of the impact of higher order modes and waveform systematics

3.1 Preamble

This chapter is a collection of different results and papers I have contributed to that investigates the impact HOMs or waveform systematics. While these were not projects I lead, my intellectual contribution was an important ingredient in allowing them to move forward. This chapter is summarized as the following:

- In Section 3.2, we show waveform systematics can play a significant role in two common precessing models. We demonstrate these systematic disagreements impact the ability to acquire accurate inferences about binary parameters. We confirm these errors to be systematic by comparing to NR waveforms. The work for this section was published in [2].

- In [3], we developed NR simulations based on the inferred parameters from analyses using semi-analytical models. In Section 3.3 (which is work published in [3]), we show via mismatch calculations that waveform systematics exist between the semi-analytical models and the newly developed NR simulations. We demonstrate through likelihood calculations that the peak likelihoods for the NR simulations and semi-analytical models

lie in different parts of parameter space.

- In Section 3.4, we show the impact of HOMs through parameter inference of a plethora of synthetic sources. Through synthetic examples, we not only show that inclusion of HOMs significantly deviates from the non-HOM result but also that the HOM robustly recovers the source parameters. The work for this section was published in [4].

- In Section 3.5, we assess the impact of HOMs on the loudest trigger, 170502, from LIGO's 2015-2017 data. While the trigger has marginal significance, we found a noticeable difference between the HOM and non-HOM analyses. The work for this section has been accepted to Astrophysical Journal[5].

3.2 Systematic challenges for future gravitational wave measurements of precessing binary black holes

In this section, we demonstrate by example several systematic issues which can complicate the interpretation of rapidly-spinning and precessing binaries. First, we provide one of the first systematic head-to-head comparisons of these models for precessing, coalescing binaries, using physically equivalent parameters for both waveforms; see also [58, 36]. We show that the two models disagree frequently for precessing systems, including parameters within the posterior distributions of gravitational wave events like GW151226 and GW170104. We focus on these two events since their posteriors are considerably different from one another, but each has similarities with other detected BBH mergers; GW151226 is similar to GW170608, and GW170104 is broadly similar to both GW170814 and GW150914. Our study differs from several previous investigations of waveform fidelity [124, 125, 110] by focusing on precessing systems and observationally-motivated parameters. The two models principally disagree when the spins are both large and precessing. GW measurements like LIGO's have not strongly prescribed whether such strongly-precessing systems are consistent with any individual observation. Using concrete examples, we remind the reader that the posterior distributions for BH spins can depend significantly on the assumed prior distributions, particularly since these distributions

ID	Model/Numerical Relativity	q	M ($M_\odot$)	χ_{1x}	χ_{1y}	χ_{1z}	χ_{2x}	χ_{2y}	χ_{2z}
A1	SEOBNRv3	1.91	60.0	-0.390	0.552	-0.346	0.174	-0.079	-0.052
A2	SEOBNRv3	3.01	26.5	0.951	-0.115	0.124	0.510	0.298	0.760
SXS:BBH:0165	Numerical Relativity	6.00	80.0	-0.058	0.776	-0.470	0.076	-0.172	-0.234
SXS:BBH:0112	Numerical Relativity	5.00	80.0	0	0	0	0	0	0

Table 3.1: **Parameters of synthetic sources**: This table shows the parameters of all the synthetic sources (waveform approximant models and numerical relativity) used in this paper. q is the mass ratio defined with $q > 1$, M is the total mass, and χ_* are the components of the normalized spins.

are often broad and nongaussian [126, 127, 128]. One astrophysically plausible prior distribution is significant BH natal spin (e.g., as motivated by some X-ray observations) and random BH spin-orbit alignment (e.g., as implied by dynamical formation scenarios). We show that, if these prior assumptions are adopted, the posterior distribution is dominated by parameters for which the models disagree even more frequently.

We perform parameter estimation on synthetic signals to demonstrate quantitatively that these disagreements lead to biases in, and different conclusions about, astrophysically relevant quantities. These synthetic signals have parameters and detector configurations consistent with observed events. Extending the study of [129], which focused on weakly precessing systems, we show that inferences about GW sources derived using the conventional configuration can frequently be biased, particularly in certain regions of the parameter space and about observationally-relevant pairs of parameters. We show that the conclusions reached can be strongly dependent on the model used. We point out that extensive followup studies – using multiple models and numerical relativity – were performed on GW150914 [120, 9, 129] and GW170104 [130, 3], producing good agreement across multiple independent calculations.

In Section 3.2.1 we compare the predictions of two models for the radiation emitted by gravitational waves from precessing BBHs. To make our discussion extremely concrete and observationally relevant, we perform these comparisons on parameters drawn from LIGO's inferences about GW151226, and from our inferences about synthetic events designed to mimic GW170104 and GW151226. Under the conventional assumptions used in this analysis, we find the two models disagree, principally when the inferred binary parameters involve large precessing spins. Because the relative probability of large and precessing spins depends on our

prior assumptions, we then repeat these comparison again, adopting prior assumptions that do not disfavor BBHs with two significant, precessing spins. To illustrate the implications of these disagreements, in Section 3.2.2, we perform several proof-of-concept parameter inference calculations using synthetic gravitational wave signals. Again, using parameters consistent with real observations (i.e., drawn from observed posterior distributions of observed BBH mergers), we show that parameter inferences performed with the two models can disagree substantially about astrophysically relevant correlated parameters, like the mass and spin of the most massive BH. To highlight the fact that these disagreements occur frequently, not merely for systems viewed in rare edge-on lines of sight, we choose synthetic binaries which are inclined by $\pi/4$ to the line of sight. In Section 3.2.3, we discuss how our results extend the broadening appreciation of potential sources of systematic error in gravitational wave measurements.

3.2.1 Models for compact binary coalescence disagree

3.2.1.1 Models for radiation from binary black holes

When inferring properties of coalescing BBHs [120, 121, 10, 130], LIGO has so far favored two approximate models for their gravitational radiation: an EOB model, denoted SEOBNRv3 [39, 106], and a phenomenological frequency-domain inspiral and merger model, denoted IMRPhenomPv2 [46], see Section 1.2 for more details.

We make use of the `lalsimulation` implementations of these two approximations, provided and maintained by their authors in the same form as used in LIGO's O1 and O2 investigations.

The coalescence time and orientation (i.e., Euler angles) of a binary are irrelevant for the inference of intrinsic parameters from gravitational wave data. As a result, and following custom in stationary-phase calculations, the IMRPhenomPv2 model does not calibrate the reference phases and time. This makes easy head-to-head comparison with time-domain calculations somewhat more difficult. Specifically, two different sets of parameters are needed to generate the same gravitational radiation in SEOBNRv3 and IMRPhenomPv2, connected by (a) a change in the overall orbital phase $\phi_{\rm orb}$; (b) a change in the precession phase of

the orbital angular momentum $\phi_{\mathbf{JL}}$; and (c) a change in the overall coalescence time t. In the approximations adopted by IMRPhenomPv2, these time and phase shifts do not qualitatively change the underlying binary or its overall orientation-dependent emission, just our perspective on it.

3.2.1.2 Binary black hole observations and model-based inference

BBH parameters are inferred through the use of Bayesian analysis with standard Monte Carlo techniques; see, e.g., [120, 23] and references therein.

For any BBH event, fully characterized by parameters x, we can compute the (Gaussian) likelihood function $p(d|x)$ for detector network data d containing a signal by using waveform models and an estimate of the (approximately Gaussian) detector noise on short timescales (see, e.g., [23, 120, 9] and references therein). In this expression x is shorthand for the set of 15 parameters needed to fully specify a quasicircular BBH. The posterior probability distribution is therefore $p(x|d) \propto p(d|x)p(x)$, where $p(x)$ is the prior probability of finding a BBH merger with different masses, spins, and orientations somewhere in the universe. LIGO–Virgo analyses have adopted a fiducial prior $p_{\rm ref}(x)$ that is uniform in orientation, in comoving volume, in mass, in spin direction (on the sphere), and, importantly for us, in spin magnitude [23, 120].

Using standard Bayesian tools [120, 23], one can produce a sequence of independent, identically distributed samples $x_{n,s}$ ($s = 1, 2, \ldots, S$) from the posterior distribution $p(x|d)$ for each event n; that is, each $x_{n,s}$ is drawn from a distribution proportional to $p(d_n|x_n)p_{\rm ref}(x_n)$. Typical calculations of this type provide $\lesssim 10^4$ samples [120, 23] from which the posterior probability distribution is inferred.

Quite clearly the choice of prior $p_{\rm ref}$ directly influences the posterior, most significantly for parameters not well constrained by the data (e.g., due to weak dependence or strong degeneracies). As a concrete example, in the left panel of Figure 3.1 we show the cumulative distribution of $\chi_{1,z}$, the component of the primary BH's dimensionless spin in the z-direction, for a synthetic source similar to GW151226. The black curve corresponds to results evaluated using the fiducial prior, where χ_1, χ_2 are distributed independently and uniformly. The red

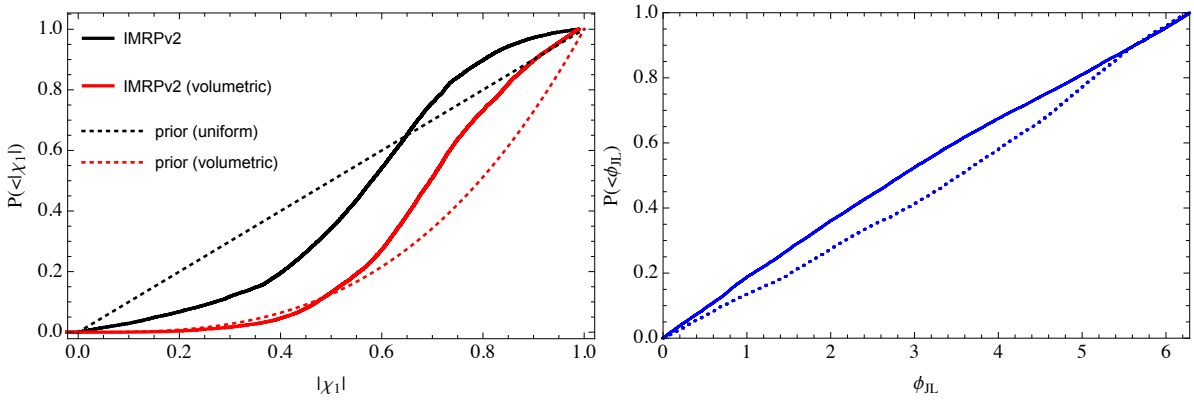

Figure 3.1: **Priors and the relative significance of large spins** *Left panel*: For a synthetic GW151226-like event (A2), the inferred cumulative posterior distribution for $|\chi_1|$ using a prior $P(<|\chi_i|) = \chi_i$ (black) and $P(<|\chi_i|) = \chi_i^3$ (red), for $i = 1, 2$. For comparison, the two priors are indicated with dotted black and red lines. The posterior probability that this synthetic event has two significant, precessing spins depends on the prior. *Right panel*: Inferred cumulative posterior distribution for $\phi_{\mathbf{JL}}$, the polar angle of $\mathbf{L}$ relative to $\mathbf{J}$, for the volumetric prior $P(<|\chi_i|) = \chi_i^3$. The solid blue line shows the results of repeating a full PE calculation, including the modified prior. The dotted blue line shows the estimated distribution calculated by weighting the posterior samples. This synthetic event was generated with parameters similar to GW151226 and analyzed with a PSD appropriate to GW150914, generated in the manner of [9].

curve is computed by drawing χ_1 from the cumulative distribution $P(<\chi_1) = \chi_1^3$, and similarly for χ_2. Henceforth we denote this the *volumetric (spin)* prior.

In the context of systematic errors and astrophysical measurements of BBHs, the choice of prior is important. Within the context of a specific astrophysical scenario or question of interest, a prior favoring large spins (or significant precession) can be appropriate. As we show later, these changes in prior can significantly increase the posterior probability of the region where model disagreement is substantial (e.g., large transverse spins, high mass ratio, and long signals).

When assessing the impact of modified priors, we exercise an abundance of caution and replicate the Bayesian inference calculations in full. In principle, with sufficiently many samples, we could estimate the posterior distribution for any prior $p(x)$ by using weighted samples. For example, we could estimate $P(<X)$ according to the modified prior $p(x)$ via the weighted empirical cumulative distribution $\hat{P}(<X) = \sum_k \Theta(X - X_k) p(x_k)/(N p_{\rm ref}(x_k))$. The approach

of reweighted posterior samples is widely proposed in hierarchical model selection [131, 132]. In practice, however, this method is reliable if and only if x_k cover the parameter space completely and sufficiently densely. In our specific circumstances, the fiducial prior $p_{\rm ref}(x)$ associates substantial prior weight near $\chi_1, \chi_2 \simeq 0$ and little probability to configurations with two large spins. As a result, rescaling from the fiducial to the volumetric prior can introduce biases into astrophysical conclusions. As a concrete example, the right panel of Figure 3.1 shows the cumulative distribution of $\phi_{\mathbf{JL}}$, the polar angle of $\mathbf{L}$ relative to $\mathbf{J}$. The solid line shows the result of a full calculation with the volumetric prior. The dotted line shows the result derived using reweighted posterior samples, starting from the fiducial uniform-magnitude prior. While the two distributions are approximately consistent in extent, the two disagree in details. If used uncritically in (hierarchical) model selection, reweighted posterior samples could lead to biased conclusions about model inference, and (in the context of our study) to biased conclusions about the relative impact of model-model systematics. Of course, a careful treatment of reweighted posterior systematics would identify this potential problem, and the need for more samples to insure a reliable answer in any reweighted application (i.e., the expected variance of the Monte Carlo integral estimate for $\hat{P}$ is large, because $p/p_{\rm ref}$ is often large).

3.2.1.3 Model-model comparisons

To quantify the difference between two predicted gravitational waves from the same binary with the same spacetime coordinates and location, we use a standard data-analysis-motivated figure of merit: the mismatch. Like other figures of merit, the mismatch is calculated using an inner product between two (generally complex-valued) timeseries $a(t), b(t)$:

$$\langle a|b\rangle = 2\int_{|f|\geq f_{\rm min}} \frac{\tilde{a}^*(f)\,\tilde{b}(f)}{S_n(|f|)} df\,, \tag{3.2.1}$$

where $S_n(|f|)$ is the noise PSD of a fiducial detector, $f_{\rm min}$ is a chosen lower frequency cut-off (typically a few tens of Hz), and the integral includes both positive and negative frequencies. Usually these comparisons also involve parameterized signals $a(\lambda, \theta)$ and $b(\lambda', \theta')$, with

maximization of the (normalized) inner product between a, b over some set of parameters Θ:

$$\langle a|b\rangle_{\Theta} = \max_{\Theta} \frac{\mathrm{Re}\,\langle a(\theta)|b(\theta')\rangle}{\sqrt{\langle a(\theta)|a(\theta)\rangle\,\langle b(\theta')|b(\theta')\rangle}} \tag{3.2.2}$$

where Θ denotes the names of the parameters in θ over which we maximize. Maximization is asymmetric; we change the parameters of only one of the two signals, effectively considering the other as "the source". When the signals a and b are real-valued single-detector response timeseries and when Θ includes only t and $\phi_{\rm orb}$, this expression is known as the *match*.

In our comparisons, we fix one of the two timeseries a generated by model A, as if it was some known detector response (e.g., from another model's prediction). The other timeseries is a predicted single-detector response $b = \mathrm{Re}F^*h$ where F is a complex-valued antenna response function and h is the gravitational wave strain. Ideally, we should evaluate b using model B and precisely the same intrinsic and extrinsic parameters, calculating the *faithfulness* [133]. The precessing models considered in this work have different time and phase conventions. In order to specify the astrophysically equivalent binary to some configuration as evolved by SEOBNRv3, we need to adopt different t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$. Reconciling the phase conventions adopted by these models is far beyond the scope of this work. However, we can find the most optimistic possible answer by maximizing the inner product over t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$, using differential evolution [134] to evolve towards the best-fitting signal. In other words, we use a figure of merit

$$\langle a|b\rangle_{t,\phi_{\rm orb},\phi_{\mathbf{JL}}} \ . \tag{3.2.3}$$

Note that, since our gravitational wave signals include higher modes, we do not simply maximize the match over $\phi_{\mathbf{JL}}$, as this would neglect the contibution to the inner product from these higher modes.

3.2.1.4 Comparison on posterior distributions

To investigate systematic errors in observationally relevant regions of parameter space,we perform model-model comparisons using samples drawn from the posterior parameter dis-

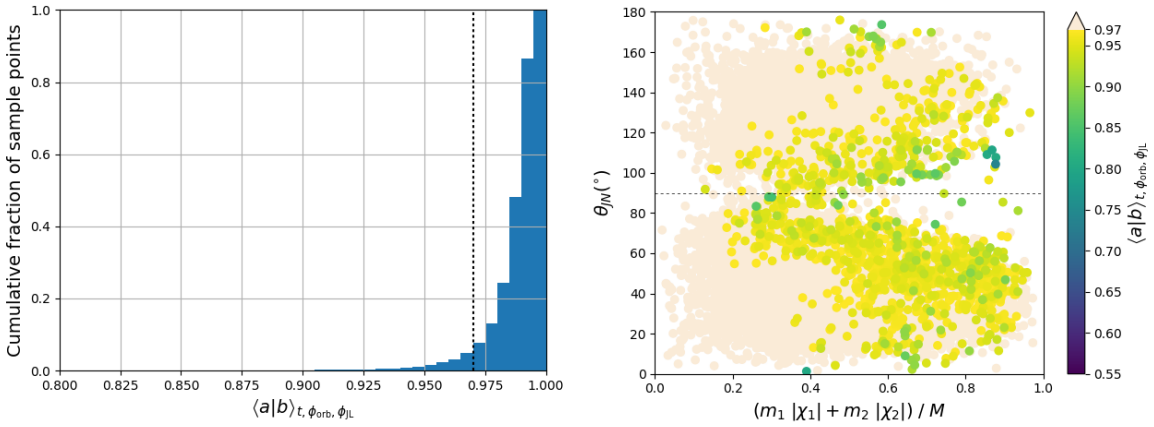

Figure 3.2: **Model-model comparison on our synthetic GW170104-like event**: Using posterior samples from our synthetic GW170104-like event (A1), we calculate model-model inner products between IMRPhenomPv2 and SEOBNRv3 waveforms, maximized over t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$. This analysis adopts the fiducial (uniform) prior on spin magnitude. In the left panel is a cumulative histogram of the maximized inner products. In the right panel the posterior samples are plotted in terms of $\theta_{\mathbf{JN}}$, the inclination of the observer relative to the total angular momentum, and a measure of the net binary BH spin. The color scale indicates the maximized inner product, with the lowest values occurring for large binary spins and preferentially near the orbital plane. The noise curve used for these calculations was the same as used in Figure 3.1.

tributions for several of LIGO's detections to date, as well as for synthetic sources. Unless otherwise noted, these comparisons are performed on the expected Hanford detector response.

Figure 3.2 illustrates our comparisons for our synthetic GW170104-like event (A1 in Table 3.1), using the fiducial spin prior. For the short waveforms needed to explain this signal, disagreement between the models is primarily associated with higher levels of precession, viewed in an orientation near the orbital plane where the effects of precession dominate.

Since lower mass systems take longer to evolve from some lower frequency to the merger, waveforms drawn from the posterior of GW151226 are significantly longer in duration. The two waveform models have significantly greater opportunity to dephase, leading to lower inner products. Figure 3.3 shows the distribution of these mismatches. The disagreement is significant over a larger portion of the parameter space than for our GW170104-like event, and is less strongly correlated with the orbital inclination $\theta_{\mathbf{JN}}$. It is clear that the spins play a leading role in producing these differences.

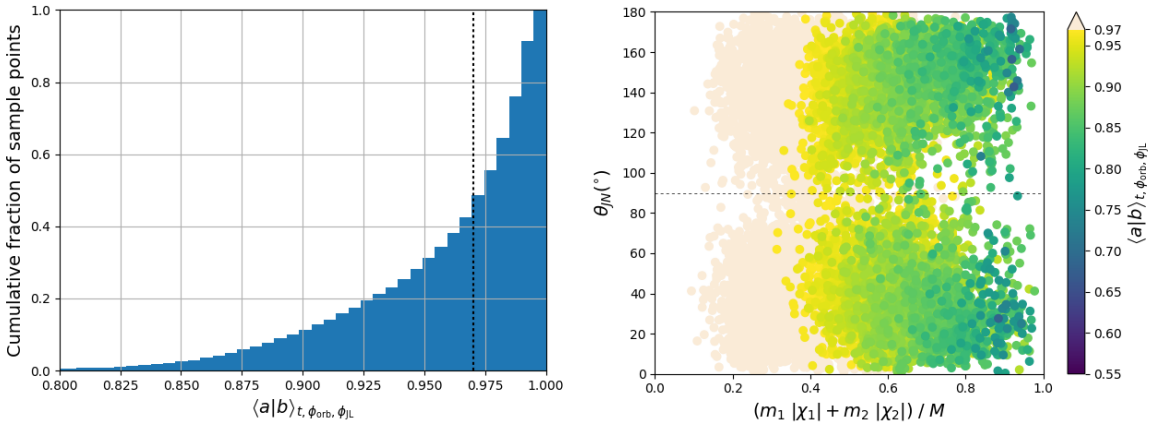

Figure 3.3: **Model-model comparison on GW151226**: As Figure 3.2 but for GW151226. The intrinsic and extrinsic parameters used in this comparison are from LIGO's O1 posterior distribution. Frequent and significant disagreement is apparent. IMRPhenomPv2 produces waveforms that are somewhat longer than SEOBNRv3 for these modest masses, leading to dephasing due to a slight difference in the rate of frequency evolution integrating over such long waveforms. This effect correlates strongly with the binary spin. The noise curve used for these calculations was calculated from data near the time to GW151226.

In Figure 3.4 we show two results for a synthetic GW151226-like event (A2 in Table 3.1), one that adopts a uniform spin prior, and the other the volumetric spin prior. As in Figure 3.3 we see that more moderate-to-highly spinning systems in the posteriors show a greater degree of diagreement between the models. The volumetric spin prior increases the support for large spins, and so increases the proportion of the posterior where model disagreement is significant.

3.2.2 Examples of biased inference of BH parameters

To illustrate the discrepancies in inferred parameters which such disagreements can cause, we select points with significant differences and generate the associated waveforms with one model, before running the full parameter estimation analysis on these waveforms using the other model. We do not add any simulated instrumental noise to the model signal in this process.

In these demonstrations of the practical differences between models from each other and from numerical relativity, we use the same parameter estimation techniques and models ap-

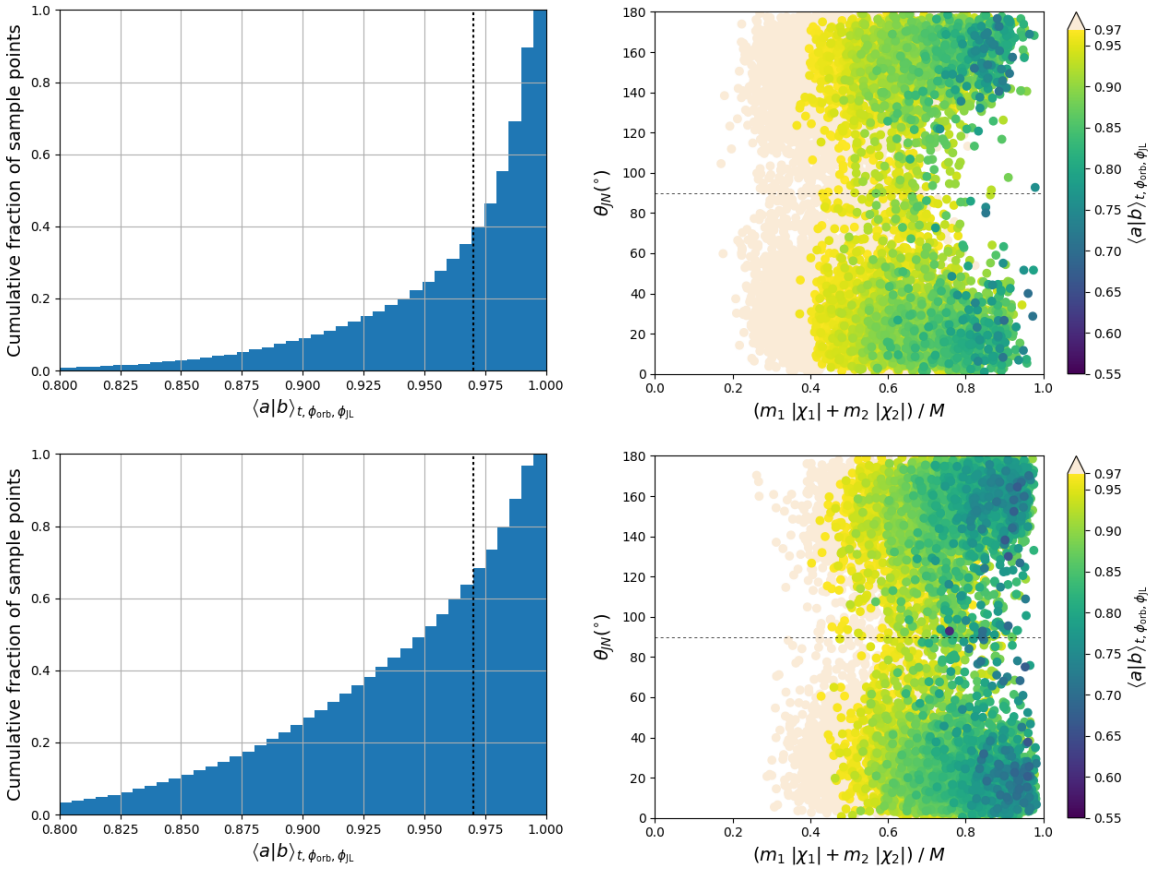

Figure 3.4: **Model-model comparison on a synthetic GW151226-like event**: As Figures 3.2 and 3.3 but for a synthetic GW151226-like event (A2). As in Figure 3.2, the top and bottom panels show the results assuming a uniform and volumetric spin prior, respectively. Adopting a volumetric spin prior noticably increases the posterior support for large spins and hence the fraction of the posterior associated with parameters where the two models disagree significantly.

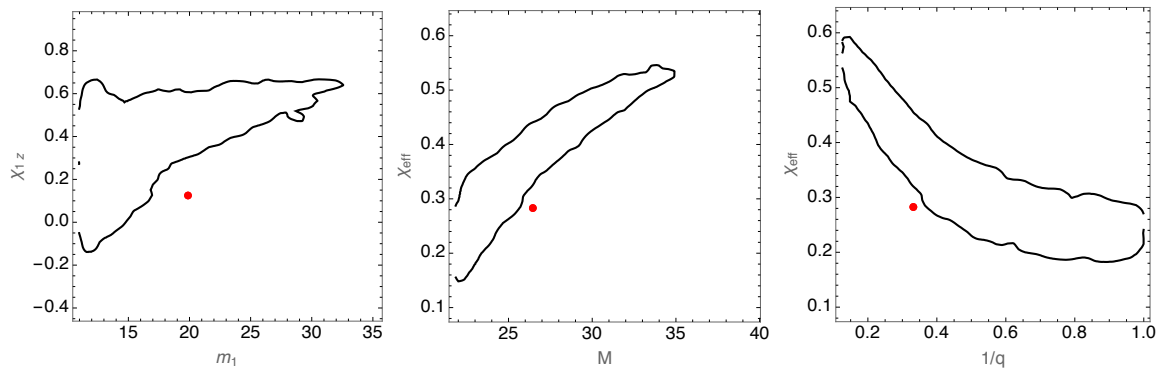

Figure 3.5: **Biased parameter recovery with IMRPhenomPv2 I: SEOBNRv3 source (A2)**: Red dot shows the parameters of a synthetic coalescing binary, whose radiation is modeled with SEOBNRv3. Binary parameters are drawn from the posterior distribution of GW151226, and are summarized in Table 3.1 as source A2. The inclination of the orbital angular momentum relative to our line of sight is $\theta_{\mathbf{JN}} = 2.48$. No synthetic noise is added to the signal. For this source, the match between the detector response predicted using IMRPhenomPv2 and SEOBNRv3 is 0.817 in Hanford, after maximizing in $t, \phi_{orb}, \phi_{\mathbf{JL}}$. Black curve shows the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomPv2 approximation. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [10, 9], using frequencies above 20Hz.

plied by LIGO to infer the parameters of the first two observed BBHs [23, 120, 10]. Figures 3.5 and 3.6 show concrete examples of biased parameter inference. In Figure 3.5, the red dot shows the parameters of a synthetic signal, generated with SEOBNRv3 using intrinsic and extrinsic parameters corresponding to a sample point in the posterior distribution for GW151226 that showed significant mismatch between waveform models. The binary parameters chosen correspond to a configuration where the models disagree (i.e., low inner product); see Figure 3.3. Our synthetic data contains only the expected detector response (the "zero noise" realization), which we interpret in the context of synthetic off-source noise with the observed frequency-dependent form.[1] The black curves show the 90% posterior confidence intervals, derived using the LALINFERENCE parameter inference engine. In Figure 3.6, we generate a synthetic source signal from a numerical relativity simulation produced by the SXS collaboration [74], using an extension to LIGO's infrastructure to designed for this purpose [140, 141]. This figure

[1] A commonly-used technique to investigate the implications of parameter inference [135, 136, 137, 138, 139, 95], the use of the "zero noise" realization allows us to compute trivially-reproducible posteriors which are centered on the true parameters yet also have the same structure (e.g., width and correlations) as would be expected from any realization of detector noise.

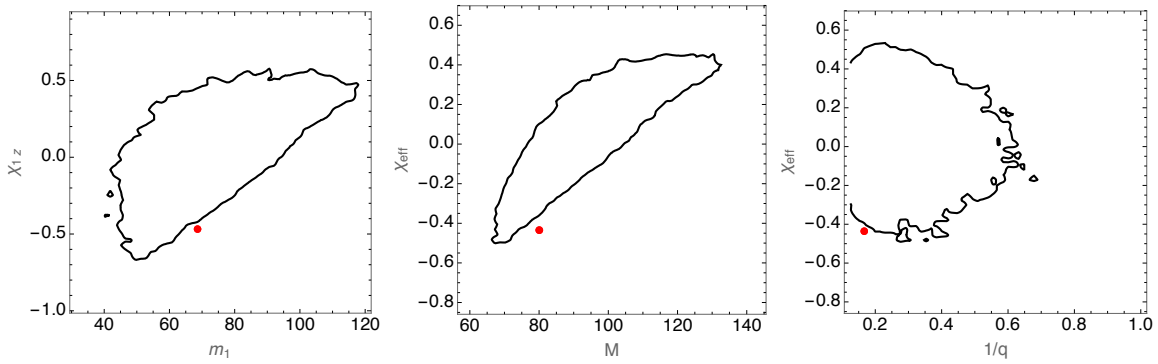

Figure 3.6: **Biased parameter recovery with IMRPhenomPv2 II: NR source**: Red dot shows the parameters of a synthetic coalescing binary, whose radiation is modeled with a numerical relativity simulation SXS BBH:0165. All simulated modes $\ell \leq 8$ are included in our synthetic signal. The detector response is calculated assuming a signal at angle $\theta_{\mathbf{JN}} = \pi/4$, at a distance so the network SNR is $\sim$10. No synthetic noise is added. The black curves show the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomPv2 approximation. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [9]. Because the $(2, 2)$ mode of this source starts at 27Hz, we only use frequencies greater than 30Hz in our analysis.

demonstrates by concrete example that the two models' disagreement can propagate into biased inference about astrophysically important binary parameters, even now in a regime of low signal amplitude and large statistical error.

3.2.3 Discussion

3.2.3.1 Mismatch does not imply bias: Examples with high mass ratio and zero spin

Due to their neglect of higher-order modes, the two models disagree significantly with numerical relativity at high mass ratio, even in the absence of spin. Several previous studies have demonstrated these modes have a significant impact on the match [142, 143, 144, 145]. For example, Figure 3.7 illustrates the mismatch introduced due to the neglect of higher order modes for non-spinning systems of varying mass ratios, with total masses $M = 80M_{\odot}$ and inclinations $\theta_{\mathbf{JN}} = \pi/4$.

A large mismatch, however, does not imply a large *bias*. As an example, Figure 3.8 shows

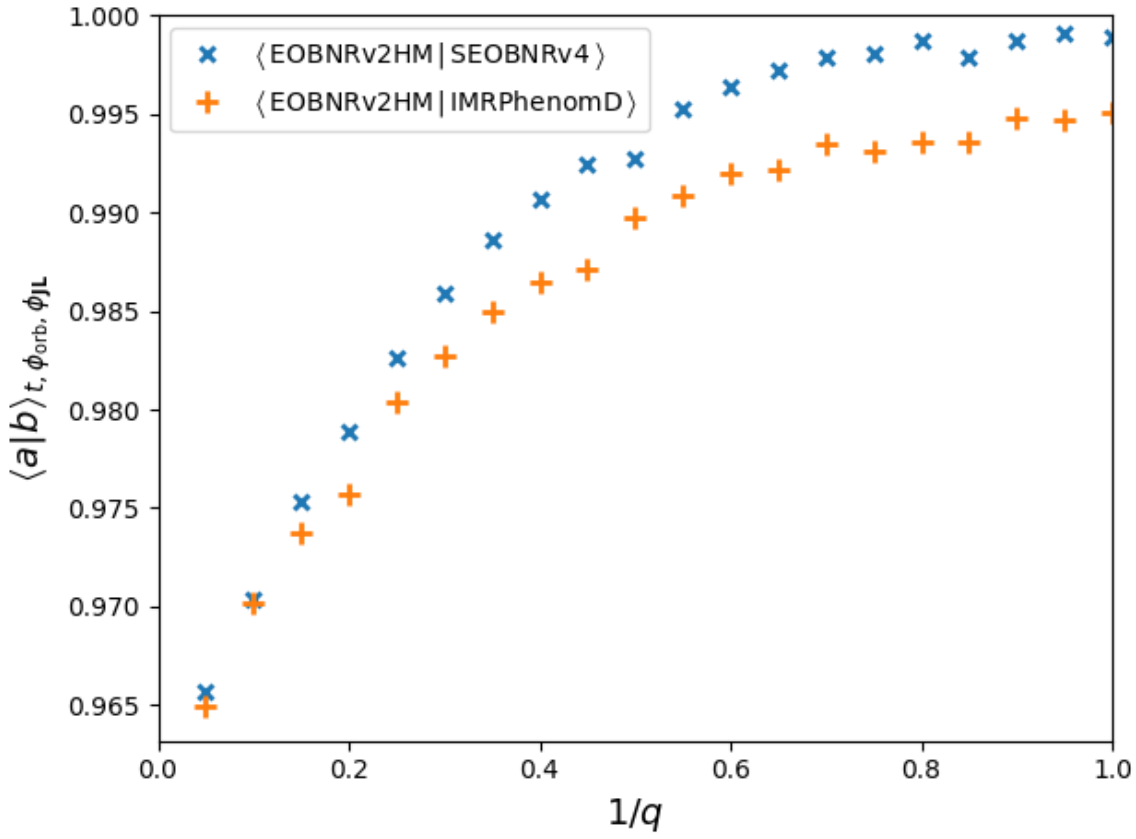

Figure 3.7: **The effect on inner product due to neglecting higher modes**: Here we generate a series of non-spinning waveforms with $M = 80M_{\odot}$ and $\theta_{\mathbf{JN}} = \pi/4$ using an EOB model that includes higher modes, EOBNRv2HM, then use the same parameters to generate waveforms with two models that do not include these higher modes, one EOB and one phenomenological – SEOBNRv4 [11] and IMRPhenomD [12]. Again we calculate the inner product maximized over $\phi_{\rm orb}$, $\phi_{\mathbf{JL}}$, and t. As higher modes are most important for heavier and unequal mass binaries, these large mismatches may be responsible for disagreements seen in Figure 3.6. Conversely, higher modes are not significant for and not included in models compared Figures 3.3 and 3.5, so are unlikely to be responsible for the large discrepancies seen there.

two sets of inferences for a signal generated from a numerical relativity simulation. This had significant mismatch with both models used to perform these inferences when evaluated at the exact binary parameters of the simulation, except t, $\phi_{\rm orb}$, and $\phi_{\mathbf{JL}}$, which again are maximized over. Despite this discrepancy, the two models draw qualitatively similar conclusions.

Another reason for mismatch that may not result in large bias is the presence of higher modes. Models which omit or include higher harmonics can often disagree substantially when judged by mismatch. In the limit of a long signal, the different harmonics have distinct time-frequency trajectories and transfer information with minimal cross-contamination [146, 147, 148]. As a result, an analysis using only *one* mode will find similar best-fitting parameters to one using multiple higher modes, but with a wider posterior due to neglected information. At higher mass and near the end of the merger, however, multiple modes are both significant and, due to their brevity, harder to distinguish. Using a simple matched-based analysis applied to hybridized nonprecessing multimodal numerical relativity simulations, [144] argued that for moderate-mass binaries, inferences based on the leading-order quadrupolar model alone would not be significantly biased, compared to the (large) statistical error expected at modest SNR; see the right panel of their Figure 1. For nonprecessing zero-spin binaries, we confirm by example that inferences about the binary are not biased. Figure 3.9 shows the posterior distributions inferred using two EOB models, one including higher-order modes (EOBNRv2HM), and the other omitting them (SEOBNRv4). The synthesised signal is a nonprecessing binary with $q = 5$ and $M = 80M_{\odot}$, generated via numerical relativity (i.e., a signal including higher order modes). Due to model limitations, these inferences are performed assuming both BHs have zero spin. This figure shows that both sets of parameter inferences are consistent with the true binary parameters used, and that inferences constructed with higher modes (via EOBNRv2HM) are both sharper and less biased than inferences that omit higher modes (via SEOBNRv4).

A large mismatch does imply, however, that the analysis is not using all available information. For example, searches for gravitational waves which neglect higher modes cannot fully capture all available signal power and a priori are somewhat less sensitive [149, 144, 143, 145];

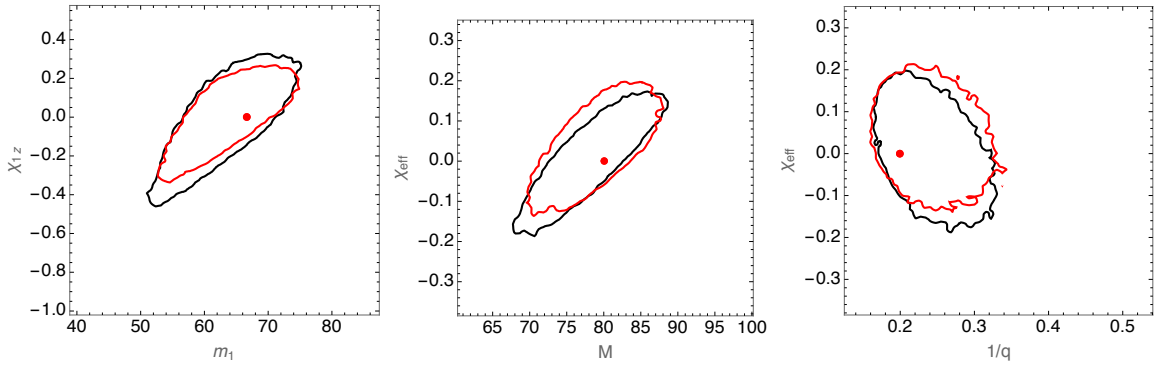

Figure 3.8: **Omitting higher modes: Unbiased parameter inference, despite a high mismatch**: Red dot shows the parameters of a synthetic nonprecessing binary, whose radiation is modeled with a numerical relativity simulation SXS BBH:0112. All simulated modes $\ell \leq 8$ are included in our synthetic signal. The detector response is calculated assuming a source with total mass $80M_\odot$ oriented at angle $\theta_{\mathbf{JN}} = \pi/4$, at a distance so the network SNR is 20. No synthetic noise is added. For this source, the best match with the IMRPhenomPv2 and SEOBNRv3 approximations is $\simeq 0.96$. The black and red curves shows the 90% confidence interval derived from a detailed parameter inference calculation using the IMRPhenomD and SEOBNRv4 approximations, respectively. Calculations are performed using a network of detectors whose noise power spectra are identical to the estimates derived for GW150914 [9].

but cf. [150]. Parameter inference calculations that use higher modes are well-known to be more discriminating about binary parameters [151, 152, 153, 154, 155, 155, 135, 136, 95]. Even for short signals associated with heavy BBHs, analyses with higher modes can draw tigher inferences about binary parameters [136, 95, 9], depending on the source; see, e.g., Figure 3.8.

3.2.3.2 Marginal distributions, degeneracy, and biases

Fortunately or not, nature and LIGO's instruments have conspired to produce short GW signals with modest amplitude to date. As illustrated by LIGO's results [10, 130] and our Figure 3.5, when using current methods (e.g., IMRPhenomPv2 and SEOBNRv3), the inferred posterior distributions for most parameters are quite broad, dominated by substantial statistical error. Inferences about individual parameters are also protected by strong degeneracies in these approximate models (e.g., in the neglect of higher-order modes) and in the physics of binary mergers (e.g., in the dependence of merger trajectories on net aligned spin). For example, the rightmost panel of Figure 3.5 shows the posterior distribution in mass ratio, q, and effective spin, $\chi_{\rm eff} = \frac{c}{G}\left(\frac{\chi_1}{m_1} + \frac{\chi_2}{m_2}\right) \cdot \frac{\hat{\mathbf{L}}}{M}$ [25, 26]; the joint posterior is tightly correlated

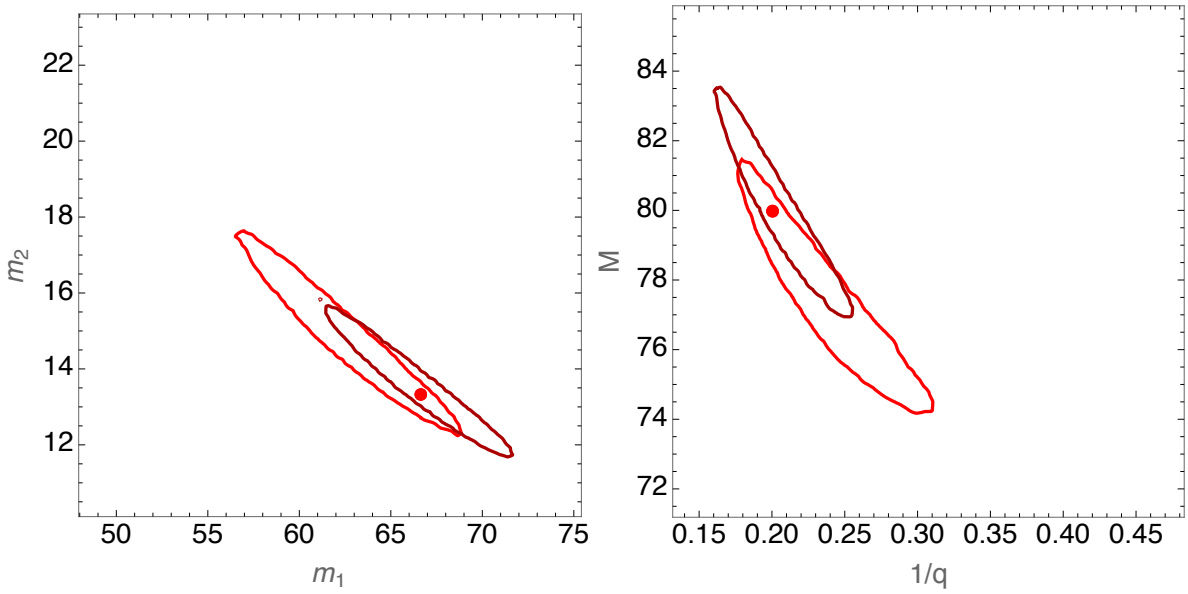

Figure 3.9: **Parameter recovery with and without higher modes (assuming zero spin)**: Red dot shows the parameters of a synthetic nonprecessing binary, generated as in Figure 3.8. The dark red contour shows inference using EOBNRv2HM (a nonspinning model including higher modes); the light red contour shows parameter inferences drawn using SEOBNRv4, assuming both BHs have zero spin. The former region is smaller than the latter, and more closely centered on the true parameters. This figure illustrates the previously-appreciated fact that inference including higher modes draws sharper conclusions with smaller biases, using the examples previously used in this work.

(and strongly biased), but the individual marginal distributions for q and $\chi_{\rm eff}$ are broad and contain the true parameters.

In principle, inference with higher modes and precession can more efficiently extract information from and produce significantly narrower posteriors for BBH mergers; see, e.g., [136, 95] as well as our Figure 3.9. Proof-of-concept new models containing these modes for precessing binaries have only recently been introduced [58, 36], and have not yet been extensively applied to parameter inference.

Run	$M_{total}/M_{\odot}$	f_{ref} [Hz]	$q = m_2/m_1$	χ_1	χ_2	GW_{Cycles}	Approximant
#1	58.49	24	0.8514	(0, 0, 0.7343)	(0, 0, -0.8278)	31.1	SEOBNRv4
#2	58.72	24	0.5246	(0.1607, -0.1023, -0.0529)	(-0.3623, 0.5679, -0.3474)	17.1	SEOBNRv3
#3	62.13	20	0.4850	(0.0835, -0.4013, -0.3036)	(-0.3813, 0.7479, -0.1021)	24.9	IMRPhenomPv2
#4	53.46	20	0.7147	(0, 0, 0.2205)	(0, 0, -0.7110)	28.2	SEOBNRv4
#5	59.11	20	0.4300	(0, 0, -0.3634)	(0, 0, -0.1256)	27.3	IMRPhenomD

Table 3.2: Numerical simulations follow-up Parameter Table (as estimated by the quoted approximant). The two runs started at 24Hz provided a fast response set of simulations while the following ones, starting at 20Hz, cover the low frequency sensitivity band. We also report the gravitational wave cycles from those frequencies to merger in the simulations.

3.3 Targeted numerical simulations of binary black holes for GW170104

This work was part of a larger project where NR follow up simulations were done based off the inference using models that was published in [3]. In this section we focus on the systematic study done comparing the NR follow up simulations and models at the same parameters.

3.3.1 Simulations of GW170104

We extracted the maximum a posteriori (MaP) parameters from (preliminary) Bayesian posterior inferences performed by the LIGO Scientific Collaboration and the Virgo Collaboration, using different waveform models [130, 23]. This point parameter estimate is one of a few well-motivated and somewhat different choices for followup parameters; however, we estimate that the specific choice we adopt will not significantly change our principal results. Table 3.2 shows parameters simulated with numerical relativity. The first two simulations have been started at a reference frequency of 24Hz (at the quoted total masses) in order to provide a fast response nonprecessing and precessing simulation to be ready to preliminarily compare with observations within 2 weeks (for the low resolution runs). The following three simulations have started from 20Hz to cover the complete nominal low frequency sensitivity band of LIGO.

Spin Conventions: $(\chi_1^x, \chi_1^y, \chi_1^z)$ are specified in a frame where (i) $\hat{L} = (0, 0, 1)$, i.e. the Newtonian orbital angular momentum is along the z-axis. (ii) the vector $\hat{n}$ pointing from m_2 to m_1 is the x-axis, (1,0,0). Note that the orientation of $\hat{n}$ is essentially undetermined by PE methods, so the choice (ii) is meant to break this degeneracy to arrive at concrete parameters.

In other words, the spin-components given below are those consistent with Eqs. (43) of Ref. [140]. For more specifics about the development of the NR simulations, see [3]

3.3.1.1 Outgoing radiation very similar for different NR methods

Following previous (targeted to GW150914) studies [13] as well as in Section 3.2, we compare the outgoing radiation mode by mode, using an observationally-driven measure: the overlap or *match*. The black and grey lines in Figures 3.10 and 3.11 show the match between the two simulations' (RIT-SXS and RIT-GT respectively) (2,2) modes, as a function of the minimum frequency used in the match. In this calculation, we use a detector noise power spectrum appropriate to GW170104, and a total mass $M_\odot$ = as given in Table 3.2. By increasing the minimum frequency, we increasingly omit the earliest times in the signal, first eliminating transient startup effects associated due to finite duration and eventually comparing principally the merger signals from the two black holes. For comparison, the red, blue, and yellow lines show the corresponding matches between RIT, SXS, and GT simulations respectively and effective one body models with identical parameters (faithfulness study). In Figure 3.10, which illustrates only nonprecessing simulations, these comparisons are made to the nonprecessing model SEOBNRv4 [11]. In Figure 3.11, which targets the two precessing UIDs, we instead compare to SEOBNRv3, which approximates some precession effects. For both nonprecessing and precessing simulations, these figures show that the different NR groups' simulations produce similar radiation, with mismatches $\leq 10^{-3}$ even at the longest durations considered. By contrast, comparisons with SEOBNRv4 and SEOBNRv3 show that these models do not replicate our simulations' results, particularly for precessing binaries. To see a further breakdown of a mode-to-mode comparison between the different simulations, see [3].

3.3.2 Likelihood of NR and models

For any proposed coalescing binary, characterized by its outgoing radiation as a function of all directions, we can compute a single quantity to assess its potential similarity to GW170104, accounting for all possible ways of orienting the source and placing it in the universe: the

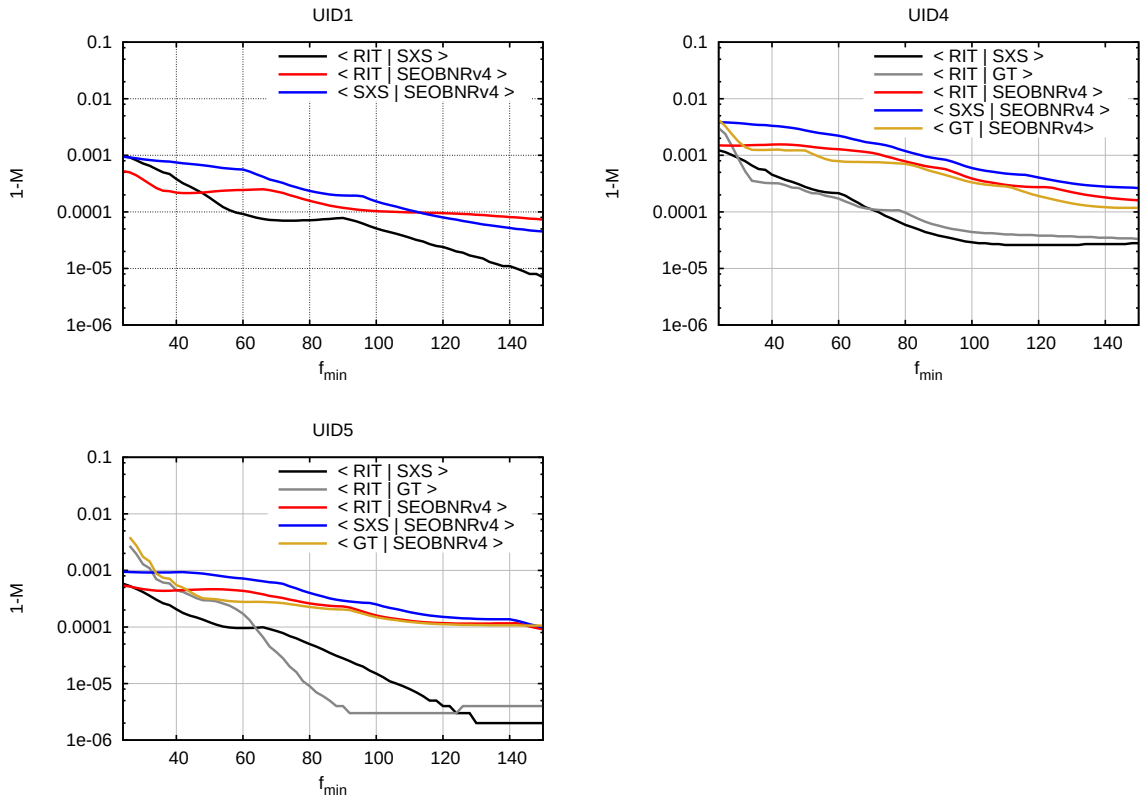

Figure 3.10: For the three nonprecessing UIDs # 1,4,5 in Table 3.2, matches between SXS, RIT, and SEOBNRv4 (2,2) modes as a function of $f_{\rm min}$, using the H1 PSD characterizing data near GW170104. We also compare with GT runs for UIDs # 4,5. Compare to also to similar plots for GW150914 [13].

UID	$\ln \mathcal{L}$(RIT)			$\ln \mathcal{L}$(SXS)	$\ln \mathcal{L}$(GT)	$\ln \mathcal{L}$(SEOB)	Model
	N100	N118	N140	L3	M120	(at NR)	
#1	60.4	61.0	61.0	60.9	-	62.7	v4
#2	61.0	60.9	60.6	60.9	-	61.4	v3
#3	60.4	60.5	60.7	60.7	-	60.4	v3
#4	60.6	60.7	60.8	60.3	60.4	62.2	v4
#5	60.0	60.0	60.1	60.0	59.8	61.2	v4

Table 3.3: **Marginalized likelihood of the data**: This table shows the results for the 5 simulations when directly compared to the data. For these results, we use the same PSD adopted in all other calculations, with $f_{\rm min} = 30$Hz (i.e. low-frequency cutoff). The first column is the UID. The second column is the estimated peak log marginalized likelihood $\ln \mathcal{L}$, maximized over binary total mass, for the NR followup simulation. The third column is the corresponding log marginalized likelihood, using exactly the same intrinsic parameters (e.g., masses and spins) as maximize the likelihood in the second column, evaluated using a phenomenological approximate model instead of numerical relativity. The fourth column is the specific model used: either SEOBNRv3 (for precessing simulations) or SEOBNRv4 (for nonprecessing simulations). To see more on this parameter estimation method, see [9, 18].

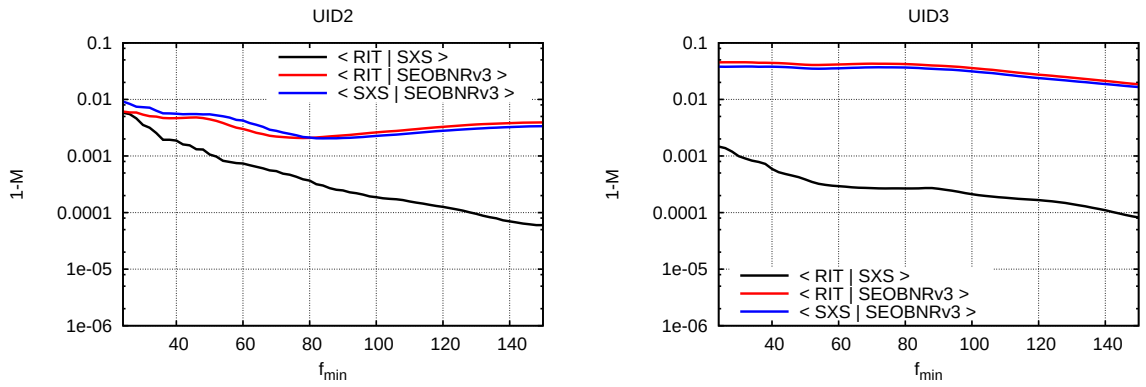

Figure 3.11: For the two precessing UIDs#2,3 in Table 3.2, matches between SXS, RIT, and SEOBNRv3 (2,2) modes as a function of $f_{\rm min}$ as a function of $f_{\rm min}$, using the H1 PSD characterizing data near GW170104. In this comparison, the $(2,2)$ mode of all three simulations and SEOBNRv3 are extracted relative to the $\hat{L}$ axis, identified from their common initial orbital parameters. While these frame identifications are coordinate-dependent for precessing binaries – implying our comparisons here could include both intrinsic disagreement and systematic error due to (say) overall misalignment – the good agreement shown in Figure 1 of [3] for the equally coordinate-dependent spins suggests that convention-dependent sources contribute little to the mismatches illustrated here.

NR Group	Label	Sim. ID	$q=\frac{m_2}{m_1}$	χ_1	χ_2	$\ln\mathcal{L}$	$\ln\mathcal{L}(SEOB)$ (at NR)	Model
RIT	a	d0_D10.52_q1.3333_a-0.25_n100	0.7500	(0, 0, 0)	(0, 0, -0.25)	63.0	62.5	v4
GT	b	(0.0,1.15)	0.8696	(0, 0, 0)	(0, 0, 0)	62.2	61.5	v4
RIT	c	q50_a0_a8_th_135_ph_30	0.5000	(0, 0, 0)	(0.490, 0.283, -0.566)	62.5	60.7	v3
BAM	d	BAM150914:24	0.8912	(-0.278, -0.605, -0.085)	(0.151, 0.396, 0.017)	62.7	61.0	v3
SXS	e	SXS:BBH:0052	0.3333	(0.001, 0.008, -0.499)	(0.494, 0.073, 0.001)	62.3	60.4	v3

Table 3.4: **Marginalized likelihood of the data: Selected other simulations**: This table shows the results for several other simulations that particularly match the data well and the SEOB model results at those parameter points. These simulations are part of the top 15 simulations in $\ln\mathcal{L}$. When comparing the NR $\ln\mathcal{L}$ values here to the ones in Table 3.3, one can see these to be generally higher i.e. better match the data. When comparing the NR $\ln\mathcal{L}$ values to the SEOB at the same points, one sees a consistent lower SEOB $\ln\mathcal{L}$ value This implies that these points were not picked for NR Followup due to the lower SEOB $\ln\mathcal{L}$ value.

marginalized likelihood ($\ln \mathcal{L}_{\rm marg}$) [22, 9, 18]. To provide a sense of scale, the distribution of $\ln \mathcal{L}_{\rm marg}$ over the posterior distribution including all intrinsic parameters is roughly universal [9], approximately distributed as $\ln \mathcal{L}_{\rm marg,max} - \chi^2/2$ where χ^2 has d degrees of freedom (i.e., a mean value of $\ln \mathcal{L}_{\rm marg,max} - d/2$, and its 90% credible interval is $\ln \mathcal{L}_{\rm marg} \geq \ln \mathcal{L} - x$, where $x = 3.89$ and $x = 6.68$ for $d = 4$ and $d = 8$, respectively). For each UID and for each proposed total mass M, direct comparison of our simulations to the data allows us to compute a single number measuring the quality of fit: the marginalized likelihood $\mathcal{L}_{\rm marg}$. The maximum value of this function (here denoted by $\mathcal{L}$) therefore measures the overall quality of fit. Table 3.3 shows $\ln \mathcal{L}$ for the five UIDs simulated here. For comparison, the last column shows $\mathcal{L}$ calculated using an approximate model for the radiation from a coalescing binary. Obviously, if these approximate models and our simulations agree, then we should find the same result for $\ln \mathcal{L}$ at the same parameters. Finally, for context, the peak value of $\ln \mathcal{L}$ computed using SEOBNRv3 with generic parameters is 63.3. If our simulation parameters are well-chosen (and if both our simulations and these models are close to true solutions of Einstein's equations), then this peak value should be in good agreement with the $\ln \mathcal{L}$ evaluated using our simulations.

First and foremost, up to Monte Carlo and fitting error, the marginalized likelihoods calculated with NR agree with each other comparing different resolutions and different approaches to solve the BBH problem, as required given the high degree of similarity between the underlying simulations. Second, the marginalized likelihoods computed at these proposed points are substantially *below* the largest $\mathcal{L}$ found with approximate models like SEOBNRv3, except for UID3. The exception here is due to the differences between the precessing models ($\ln \mathcal{L}$ was calculated with SEOB but the parameters were suggested with IMRPhenomPv2), see Appendix B from [3]. Likewise, the binary parameters at which the peak value of $\mathcal{L}$ occurs for SEOBNRv3 are substantially different from any of the proposed parameters explored here. This discrepancy suggests that the model-based procedure that we adopted to target our followup simulations was not effective at finding the most likely parameters, as measured with $\ln \mathcal{L}$. The poor performance of our targeted followup cannot simply reflect sampling error; even though the likelihood surface is nearly flat near the peak, so small errors are amplified

in parameter space, this near-flatness also insures that systematic offsets *should* produce a small change in $\ln \mathcal{L}$, *if* the underlying waveform calculations agree, see Appendix B from [3]. Instead, we suspect the biases in $\mathcal{L}$ arise because the models only approximate the correct solution of Einstein's equations. Third, we confirm our hypothesis in Table 3.4 simply by demonstrating that other simulations (not performed in followup) fit the data substantially better than our targeted parameters.

ϕ	ϕ_{merger}	$\ln \mathcal{L}$	$M_z/M_\odot$
0	0	62.3	54.9
30	19.5	62.5	55.2
60	34.8	62.2	54.1
90	56.5	62.5	54.4
120	98.5	61.6	54.1
150	146.5	60.6	54.5
210	194.7	59.3	55.1
310	294.0	60.4	54.6
A	B	C	RMS
1.23 ± 0.21	-0.75 ± 0.15	61.1 ± 0.15	0.38
A_{merger}	B_{merger}	C_{merger}	RMS_{merger}
1.08 ± 0.18	-0.47 ± 0.19	61.1 ± 0.15	0.37

Table 3.5: The log-likelihood of the NQ50TH135 series [14]. Fittings of the form $\ln \mathcal{L} = Asin(\pi/180\phi + B) + C$ is also given for both the initial ϕ and ϕ_{merger}.

On the one hand, NR followup simulations guided by the models (as displayed in Table 3.3) leads to lower marginalized likelihoods ($\ln \mathcal{L}$). Conversely, other simulations shown in Table 3.4 produce higher $\ln \mathcal{L}$, at points in parameter space where the models predict lower $\ln \mathcal{L}$. This discrepancy suggest the two processes ($\ln \mathcal{L}$ evaluated with NR and with the models) favor different regions of parameter space. In particular, Table 3.4, which has one of the largest values of $\ln \mathcal{L}$ among all of the (roughly two thousand) simulations available to us, shows that the top precessing simulation is `q50_a0_a8_th_135_ph_30`. This simulation has a mass ratio of 1:2, i.e. q=1/2, where the smaller hole is nonspinning and the larger hole is spinning with an intrinsic spin magnitude of 0.8 and pointing initially in a direction downwards with respect to the orbital angular momentum (θ=135 degrees) and an angle of 30 degrees from the line joining the two black holes (ϕ=30 degrees). This simulation belongs to a family of 6 simulations performed in Ref. [14] labeled as NQ50TH135PH[0,30,60,90,120,150]. Those

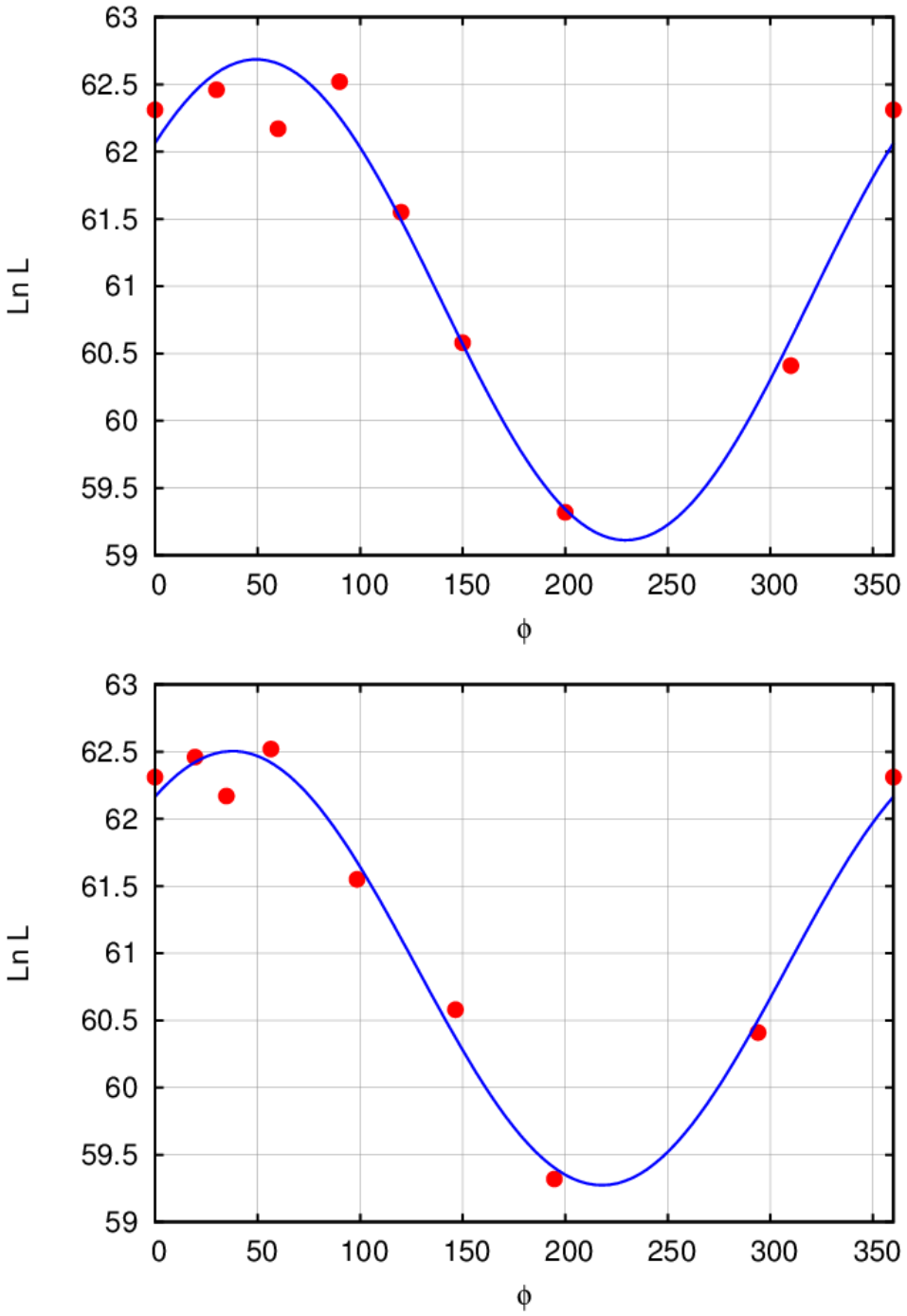

Figure 3.12: The log-likelihood of the NQ50TH135 series [14] assuming a period of 2π versus initial angle (top panel) and merger angle (bottom panel.) Data (red) and fits (blue) are given in Table 3.5.

runs, supplemented by two control runs with angles $\phi = 200$, 310 we performed for this work originally from [3] are displayed in Figure 3.12 versus the $\ln \mathcal{L}$ for this GW170104 event. The lower panels plots all those simulation with respect to their ϕ-angle at merger as defined in Ref. [14] and given in Table XXI in that paper. The continuous curve provide a fit (detailed in Table 3.5) for such values as reference and an estimate of the maximum value located near the phi=30 simulation. For a further discussion on how this sensitivity to the orientation can be used more broadly with similar simulations, see [3]

3.4 Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems

3.4.1 Gravitational Wave Model

For a discussion of the parameters that described a GW, see Section 1.1.2. As is discussed in Section 1.3.1.1, it is common practice to introduce a complex gravitational-wave strain

$$h_+(t;t_c,\iota,\phi_c,\vec{\lambda}) - \mathrm{i}h_\times(t;t_c,\iota,\phi_c,\vec{\lambda}) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell} h^{\ell m}(t-t_c;\vec{\lambda})\,{}_{-2}Y_{\ell m}(\iota,\phi_c,)\,, \tag{3.4.4}$$

which is subsequently decomposed into a basis of spin-weighted spherical harmonics ${}_{-2}Y_{\ell m}$. Here $\vec{\lambda} \equiv (q, M, \chi_{1z}, \chi_{2z})$ is used to denote the signal's dependence on the intrinsic parameters, ι is the inclination angle between the orbital angular momentum of the binary and line-of-sight to the detector, t_c is the coalescence time, and ϕ_c is the orbital phase at coalescence. Most gravitational waveform models make predictions for the modes $h^{\ell m}(t)$, from which the gravitational-wave strain detected by a ground-based interferometer,

$$h(t;\vec{\Lambda}) = \frac{1}{r}F_+\left(\mathrm{ra},\mathrm{dec},\psi\right)h_+(t;t_c,\iota,\phi_c,\vec{\lambda}) + \frac{1}{r}F_\times\left(\mathrm{ra},\mathrm{dec},\psi\right)h_\times(t;t_c,\iota,\phi_c,\vec{\lambda})\,, \tag{3.4.5}$$

is readily assembled. The signal's dependence on four additional extrinsic parameters are the polarization angle (ψ), the luminosity distance to the source's center-of-mass (r), and sky location determined by the right ascension (ra) and declination (dec). The antenna patterns $F_{(+,\times)}$ project the GW's +- and ×-polarization states, $h_{(+,\times)}$, into the detector's frame. We shall use $\vec{\Lambda} \equiv (\mathrm{ra}, \mathrm{dec}, \psi, r, t_c, \iota, \phi_c, \vec{\lambda})$ to denote the signal's dependence on all 11 parameters defining the problem.

Until recently, all spinning IMR models had set $h^{\ell m} = 0$ except for the dominant $h^{2,\pm 2}$ quadrupole modes. The expectation had been that higher modes won't substantially affect parameter inference for the O2 gravitational-wave observations, which are characterized by low SNRs and mostly face-on events of near-equal mass [156, 142, 129].

Over the past year or so, three new aligned-spin IMR models have been built to include non-quadropole modes: (i) a phenomenological frequency-domain model, IMRPhenomHM [47], includes the $(\ell, |m|) = (2,2), (3,3), (4,4), (2,1), (3,2), (4,3)$ modes; (ii) an effective-one-body time-domain model, SEOBNRv4HM [42], includes a similar set of $(\ell, |m|) = (2,2), (3,3), (4,4), (5,5), (2,1)$ modes; (iii) a time-domain surrogate model for hybridized nonprecessing numerical relativity waveforms, NRHybSur3dq8 [57], includes all of the $\ell \leq 4$ and $(5,5)$ spin-weighted spherical harmonic modes but not the $(4,1)$ or $(4,0)$ modes.

Our study will use NRHybSur3dq8 as it both includes the most modes and is expected to be more accurate when evaluated within its training region (cf. Fig 6 from Ref. [57]) of mass ratio $q \leq 8$, and $|\chi_{1z}|, |\chi_{2z}| \leq 0.8$. For the 20 Hz starting frequency considered here, this model is valid for the entire LIGO band for stellar mass binaries with total masses as low as $2.25\,M_\odot$. We evaluate the model through the Python package GWSurrogate [2] [158, 53]. The GWSurrogate package provides direct access to the GW's harmonic modes $h^{\ell m}(t)$ appearing in the sum (3.4.4).

By comparing to NR, Ref. [57] has computed the NRHybSur3dq8 model's mismatches (averaged over many points on the sky) as a function of total mass using the Advanced LIGO design sensitivity noise curve. For the $120M_\odot$ total mass systems predominantly used in

[2] We use GWsurrogate version 0.9.{4,5}, which exactly agrees with the lalsimulation [157] implementation of the NRHybSur3dq8 model.

our studies, the single-detector mismatches have a median value of 1×10^{-5}. A sufficient condition for two waveform models (in this case NR and NRHybSur3dq8) to be considered indistinguishable is [159, 160, 161, 129]

$$\mathcal{M} < \frac{D}{2\rho^2}\,, \tag{3.4.6}$$

where $\mathcal{M}$ is the mismatch and ρ is the signal-to-noise ratio (SNR). Here D is an unknown constant that is sometimes associated with the number of model parameters [162], with $D = 4$ for our spinning BBH model. Furthermore, if the likelihood can be approximated by a Gaussian then an expression for D can be obtained in terms of a chi-squared distribution with 4 degrees of freedom [163]. Using this value for D and a typical mismatch value quoted above, we find that the NRHybSur3dq8 model will give robust parameter estimates so long as $\rho \lesssim 450$. Even using pessimistic values ($D = 1$ and the 95th percentile of mismatch errors 7×10^{-5}) we find that NR and our model will be indistinguishable according to Eq. (3.4.6) so long as $\rho \lesssim 85$.

For context, we note that in the first and second observing runs most BBH signals had a network SNR of about 15 and spanning a range of 10 to 30. In the upcoming observing run we would expect typical BBH SNRs to be between 10 and roughly 40, based on the cumulative distribution of the loudest SNR ρ among n identified events ($[1 - (\rho/10)^3]^n$ using a fiducial value $n = 30$). We caution the reader that in practice the condition in Eq. (3.4.6) should only be taken as a rough estimate. For instance, it features an unknown constant D while the NR waveforms themselves have small, systematic sources of error that would prevent any model to claim indistinguishability from general relativity beyond estimates of this systematic error [78]. Finally, the definition of “indistinguishable" is not synonymous with “identical posterior distributions". Indeed, Figure 3.16 shows that even for simple systems at low SNR, which easily satisfy Eq. (3.4.6), there can be noticeable discrepancies between the recovered posteriors. For example, using a single interferometer the mismatch between $\ell_{\rm max} = 5$ and $\ell_{\rm max} = 2$ models for a non-spinning, equal-mass system is 0.0021, and so Eq. (3.4.6) is satisfied at SNRs less than 30.

Due to the absence of higher-mode models for spinning BBH systems until recently, previ-

ous parameter-inference studies that have focused on the information content available higher modes have either used quadrupole-only (recovery) models or leveraged the Fisher matrix framework. For high-accuracy, high-SNR scenarios involving the 3-detector network neither of these are fully sufficient. For example, with the quadrupole-only model the reference ("true") posterior will not be possible to compute in principle. Additionally, some of these models may have modeling errors in the dominant mode that could become noticeable at high SNR [164, 2, 165].

3.4.2 Bayesian Inference

See [22, 102, 1] papers, Chapter 2, and Section 1.3 for technical details of the PE used in this section. To achieve rapid turnaround times, we use the new GPU-accelerated implementation of ILE [102]. Working on the CARNiE cluster, which includes 15 NVIDIA Tesla V100 GPU-enabled nodes, our current configuration completes each of the binary black hole analyses presented in this work in about 15 to 20 hours. When using all 15 GPUs, a single ILE step for an SNR=30 case takes about 1 hour to finish.

Following the RIFT algorithm [1], we iteratively construct an approximation to the likelihood by generating and drawing from approximate posterior distributions, until our posterior distribution converges. At each iteration, the likelihood is approximated using Gaussian process regression with a squared-exponential kernel, with hyperparameters tuned to the likelihood evaluations available at that iteration. In this work, we present PE result from sources listed in Table 3.6. All synthetic datasets use PSDs generated from data near GW170814 [166], when all three detectors were operational, and are created with zero noise realizations. Specifically the synthetic detector data is exactly equal to the expected response due to our GW source. Since detector noise is assumed to be colored Gaussian noise with zero mean, using zero noise with the likelihood defined in Eq. 2.1.2 makes our analysis equivalent to an average over an ensemble of analyses which use infinitely many noise realizations [129]. For all runs, $f_{\rm min}$ and $f_{\rm max}$ from Eq. 1.1.9 are 20 Hz and 2000 Hz, respectively.

3.4.3 Intrinsic-parameter biases

Each synthetic dataset includes an injected signal from the expected response at each detector due to our GW source using the NRHybSur3dq8 model and including all of the surrogate's available $\ell_{\rm max} = 5$ modes (this model only include the (5,5) and excludes the (4,$\pm$1) and (4,0) modes). The model generates a waveform such that the instantaneous initial frequency of the $(2, 2)$ mode has a frequency of 8 Hz, which ensures the $(5, 5)$ mode's instantaneous initial frequency is out-of-band. We taper the beginning and end portions of the waveform to avoid artificial oscillations in the Fourier domain. In particular, since NR waveforms (and therefore the NRHybSur3dq8 model) do not go to zero by the end of the simulation, we have found it necessary to taper the last portion of the ringdown signal.

We adopt conventional mass and distance priors, uniform in detector-frame mass and in the cube of the luminosity distance. For our nonprecessing spins, we adopt a uniform prior for $\chi_{i,z} \in [-0.9, 0.9]$. Section 3.4.4.1 considers the effect of using an alternative spin prior in the context of high SNR events.

Each of the following subsections describe a set of related runs, varying one of the problems' parameters at a time. For each source configuration, we present parameter estimates recovered using all of the available higher modes $\ell_{\rm max} = 5$ (we may sometimes refer to this as the "true" or reference posterior) and compare with posteriors recovered using the same model restricted to only the $\ell_{\rm max} = 2$ modes (using $|m| = \{2, 1\}$). In Sections 3.4.3.1 ($q = 1$), 3.4.3.2 ($q = 4$), and 3.4.3.3 ($q = 7$) we vary the spin configurations of $\chi_{1\rm z} = \chi_{2\rm z} = \{-0.8, -0.5, 0.0, 0.5, 0.8\}$ while keeping the network SNR fixed at 30 [3]. For this sequence of runs, our choice of inclination angle, $\iota = 3\pi/4$, is neither face-on nor edge-on, but rather constitutes a "general" configuration. In Section 3.4.3.4 we consider varying the SNR to explore its effect on marginalized posterior distributions.

It is known that the contribution of subdominant modes towards the signal's power increases as the inclination angle is increased from a face-on ($\iota = 0$) to an edge-on ($\iota = \pi/2$)

[3] Given a fixed starting frequency, systems with their BH component spins (anti-)aligned with the orbital angular momentum will be (shorter) longer. As a result, to achieve a fixed SNR the spin (anti-)aligned systems must be place located (closer) farther as compared to a reference non-spinning system.

ID#	ι	q	M ($M_\odot$)	χ_{1z}	χ_{2z}	SNR
1	$\pi/4$	2.267	127.1	0.72	0.0	30
2	$3\pi/4$	1.00	120.0	-0.80	-0.80	30
3	$3\pi/4$	1.00	120.0	-0.50	-0.50	30
4	$3\pi/4$	1.00	120.0	0.0	0.0	10,30,70
5	$3\pi/4$	1.00	120.0	0.50	0.50	30
6	$3\pi/4$	1.00	120.0	0.80	0.80	30
7	$3\pi/4$	4.00	120.0	-0.8	-0.8	30
8	$3\pi/4$	4.00	120.0	-0.5	-0.5	10,30,70
9	$3\pi/4$	4.00	120.0	0.0	0.0	30
10	$3\pi/4$	4.00	120.0	0.5	0.5	30
11	$3\pi/4$	4.00	120.0	0.8	0.8	30
12	$3\pi/4$	7.00	120.0	-0.8	-0.8	30
13	$3\pi/4$	7.00	120.0	-0.5	-0.5	30
14	$3\pi/4$	7.00	120.0	0.0	0.0	30
15	$3\pi/4$	7.00	120.0	0.5	0.5	30
16	$3\pi/4$	7.00	120.0	0.8	0.8	30

Table 3.6: **Parameters of synthetic sources**: This table shows the parameters of all the synthetic sources used in this paper. ι is the inclination angle between the line of sight of the observer and the total angular momentum vector, q is the mass ratio defined with $q > 1$ (see Eq. 1.1.21), M is the detector-frame total mass, and χ_* are the components of the normalized spins (see Eq. 1.1.23). As we use a non-precessing model, we set all of the in-plane spin components to 0. All luminosity distances are set such that the network signal-to-noise ratio achieves the value specified under the SNR column. For example, in our $q = 7$ sequence the most extreme values of spin, $\chi_{\rm eff} = -0.8$ and $\chi_{\rm eff} = 0.8$, are located at 181.4720 Mpc and 452.5185 Mpc, respectively. This large discrepancy in distance is due to the orbital hangup effect and is explained in greater detail in Figure 11 of [4]. Other extrinsic parameters are fixed to the following values: right ascension is RA=0.0, declination is DEC=1.5707963, and the polarization angle is $\psi = \pi/4$.

configuration. As such, we expect our observed biases to be larger (smaller) when compared to a face-on (edge-on) system at the same network SNR value. This general expectation was recently confirmed by Kalaghatgi et al. [167], where the importance of subdominant modes for non-spinning systems was quantified by systematically varying the inclination angle across a range of values. In our study we have instead fixed the inclination angle to a value typical of an O2 event [15] while systematically exploring the impact due to spin. As such our results are complementary to those of Ref. [167].

3.4.3.1 q=1

We first look at a set of equal mass runs with the different spin configurations mentioned above. It is well known that the relative power of subdominant harmonic modes are minimized for equal mass BBH systems, so these cases are expected to minimize bias. Previous studies [156, 142, 144, 129] have either found negligible bias (for face-on systems), small bias (for edge-on systems), or quoted results averaged over the source orientation where again only very small biases were found. At the time of these studies [156, 142, 144, 129], however, there were no recovery models for near-equal mass spinning BBH systems including subdominant modes so these results were only suggestive. Here we confirm the general expectation of smaller bias at $q = 1$, while also making more precise the nature of the bias by comparing the true posterior to the approximate one found with $\ell_{\rm max} = 2$ modes only. For example, in all cases the true posterior's peak is located at $q = 1$, while some of the biased posteriors have a non-negligible offset often peaking closer to $q \sim 1.25$. From Figure 3.13 we also observe noticeable shifts in the posteriors 90% confidence region for anti-aligned configurations.

Figure 3.13 shows the posterior distributions of the intrinsic parameters for all the different spin configurations. The solid lines represent runs that were done with $\ell_{\rm max} = 2$ modes, and the dashed lines represent runs that include all available $\ell_{\rm max} = 5$. For each run, there is some degree of difference between the $\ell_{\rm max} = 2$ and $\ell_{\rm max} = 5$ runs. As anticipated by Ref. [144], which used a non-Bayesian approach and a single detector, this discrepancy between the two distributions become more extreme as the spins increase toward negative spin. For example, for negative spins there are noticeable shifts in the M vs $\chi_{\rm eff}$ posteriors. We emphasize that even for the simplest case (equal mass and zero spin), differences between the two results are visible. Although parameter recovery is not biased in the sense that all of the injection values lie within their 90% confidence regions, it is also clear from the figure that the median recovered using all subdominant modes is almost always closer to the injection value. This is contrary to the general expectation that subdominant modes are largely irrelevant for equal-mass systems [156, 142, 129]. Section 3.4.3.4 explores how different network SNRs affect the bias for these systems; Appendix 3.4.5 follows up on the curious differences seen in the simplest case of zero

spin, equal mass.

3.4.3.2 q=4

We next increase our set of sources to $q = 4$, a configuration that is most relevant to GW190412-like events. Similar to the $q = 1$ case, as far as we are aware, the existing literature for parameter estimation is comprised of results for non-spinning recovery models [156], results for near-equal mass without multi-mode recovery models [129], or Fisher matrix-based studies [142, 144]. None of those studies consider the 3-detector network configuration and a multi-modal recovery model with fully Bayesian inference. At larger mass ratios, our study confirms the general expectations described in Ref. [144], although the observed bias is often even larger than expected; compare to the typical errors indicated by corresponding green, red, and blue curves in Figure 6 of Ref. [144] for our fiducial mass. We also are able to more carefully quantify the nature of the bias by comparing to the true posteriors. In particular, similar to the $q = 1$ systems just considered, neglecting subdominant modes consistently shifts the posterior towards more extreme anti-aligned spin configurations with lighter total mass.

Figure 3.14 shows the posterior distributions for $\chi_{\rm eff}$ vs q and $\chi_{\rm eff}$ vs M for all the different spin configurations. The solid lines again represent runs that were done with $\ell_{\rm max} = 2$ modes, and the dashed lines represent runs that include all available $\ell_{\rm max} = 5$. Similar to Section 3.4.3.1, we again see that the differences become more extreme as the spins increase toward negative spin. Comparing the same spin configures between $q = 1$ and $q = 4$ runs, it is clear that increasing the mass ratio dramatically increases the bias between the non-HM and HM runs. In particular, there are now many cases where parameter estimates recovered with $\ell_{\rm max} = 2$ modes do not lie within their 90% confidence regions. Looking at the two-dimensional posteriors, for example, shows many cases where either the $\ell_{\rm max} = 2$ posterior either does not contain the injection value or it is noticeably shifted from the true posterior. By comparison, in almost all of the $\ell_{\rm max} = 5$ cases, the marginal posteriors almost perfectly peak at the true parameters. One notable exception is the $\chi_{\rm eff} = -0.8$ case (the purple distributions in Figure 3.14) where the true parameters seem to lie just inside the 90% confidence region. We suspect

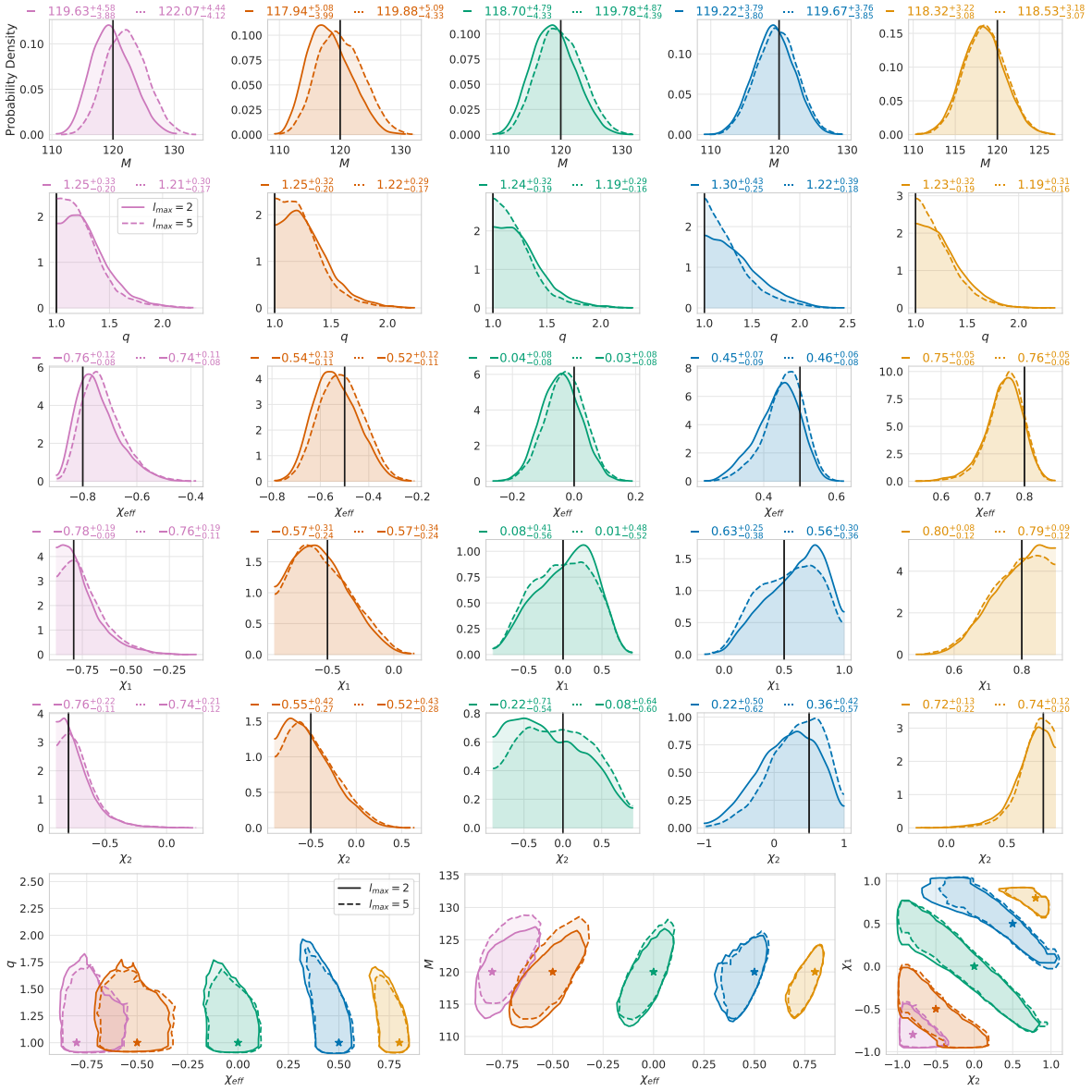

Figure 3.13: **Non-HM and HM runs for q=1 spin set, with SNR=30 and M=120**: The first five rows show the $M, q, \chi_{\rm eff}, \chi_{1z}, \chi_{2z}$ one-dimensional marginal distributions, where among this set of figures each column corresponds to a different synthetic source recovered with either all $\ell_{\rm max} = 5$ modes (dashed line) or $\ell_{\rm max} = 2$ modes (solid line). Our figures are organized such that the injected spin is systematically increased from left to right, where the synthetic source runs are ID2 ($\chi_{\rm eff} = -.8$), ID3 ($\chi_{\rm eff} = -.5$), ID4 ($\chi_{\rm eff} = 0$), ID5 ($\chi_{\rm eff} = .5$), and ID6 ($\chi_{\rm eff} = .8$). In each figure's title, we report the median value and the 90% confidence intervals of the marginalized 1D distribution for the $\ell_{\rm max} = 2$ (left) and $\ell_{\rm max} = 5$ (right) cases. A solid black vertical line denotes the true parameter value. The final bottom row corresponds to the joint distributions for q vs $\chi_{\rm eff}$, M vs $\chi_{\rm eff}$, and $\chi_{1,z}$ vs $\chi_{2,z}$ for all five injections.

this is due to a combination of (i) the injection being very close to the boundaries of the prior and (ii) the posterior for a $\chi_{\rm eff} = -0.8$ injection is much wider than the corresponding $\chi_{\rm eff} = 0.8$ value, which does not show this unexpected behavior.

3.4.3.3 q=7

Finally, we analyze sources with $q = 7$. Figure 3.15 shows the posterior distributions of the intrinsic parameters for all the different spin configurations. The solid lines again represent runs that were done with $\ell_{\rm max} = 2$ modes, and the dashed lines represent runs that include all available $\ell_{\rm max} = 5$. As expected and consistent with the trend seen in the previous two subsections, we see substantial biases are often introduced in M, q and $\chi_{\rm eff}$ if higher modes are omitted, especially for systems with large negative spin. Only the higher-mode model is able to make reliable parameter estimates, except for the large, positive spin configurations where a quadrupole-only model continues to do reasonably well. In some cases the the biased posterior doesn't even overlap with the true one, which would be problematic for likelihood-reweighting techniques [168], which require similar posterior distributions.

Somewhat unexpectedly, however, is that the $\chi_1 = \chi_2 = 0.8$ system's posterior shows almost no effect from neglecting subdominant modes; any effect that is present is smaller than the corresponding equal-mass system with $\chi_1 = \chi_2 = -0.8$. We believe this can be explained by the orbital hangup effect [169], whereby given two otherwise identical systems the one with larger aligned spin will experience more orbits before merger. Consequently, the $\chi_1 = \chi_2 = 0.8$ configuration will have more in-band cycles, and subdominant modes are known to be suppressed during the inspiral phase. We briefly elaborate on this effect in the conclusions.

3.4.3.4 Effect of Network SNR on Biases

In the previous subsections, it was shown that a significant bias exists at SNR=30, even for the simplest systems. This subsection is dedicated to investigating how the SNR affects the bias. Here we use all the different SNR runs from ID4 and ID8 in Table 3.6. Figures 3.16

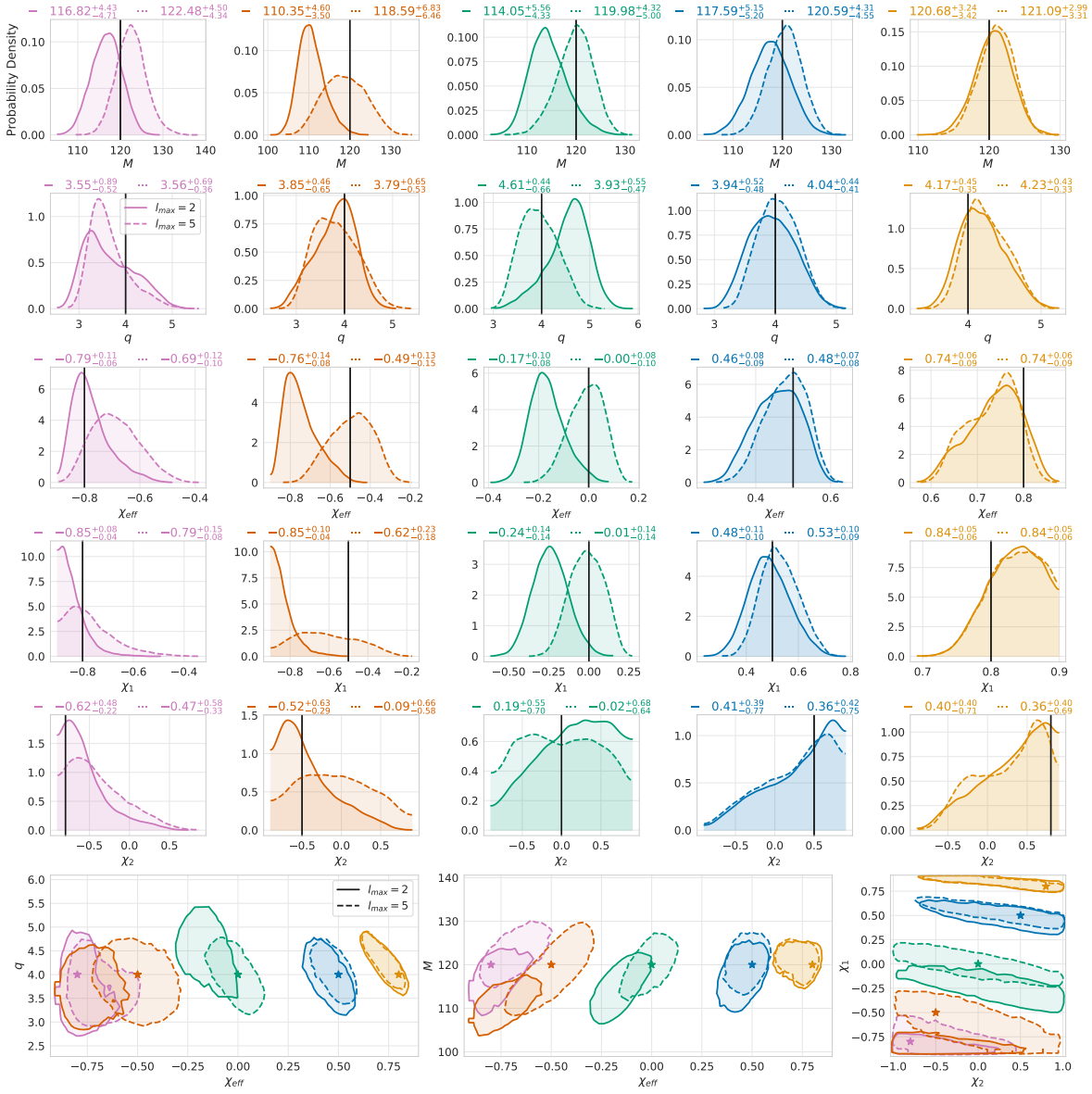

Figure 3.14: **Non-HM and HM runs for q=4 spin set, with SNR=30 and M=120**: The first five rows show the $M, q, \chi_{\rm eff}, \chi_{1z}, \chi_{2z}$ one-dimensional marginal distributions, where among this set of figures each column corresponds to a different synthetic source recovered with either all $\ell_{\rm max} = 5$ modes (dashed line) or $\ell_{\rm max} = 2$ modes (solid line). Our figures are organized such that the injected spin is systematically increased from left to right, where the synthetic source runs are ID7 ($\chi_{\rm eff} = -.8$), ID8 ($\chi_{\rm eff} = -.5$), ID9 ($\chi_{\rm eff} = 0$), ID10 ($\chi_{\rm eff} = .5$), and ID11 ($\chi_{\rm eff} = .8$). In each figure's title, we report the median value and the 90% confidence intervals of the marginalized 1D distribution for the $\ell_{\rm max} = 2$ (left) and $\ell_{\rm max} = 5$ (right) cases. A solid black vertical line denotes the true parameter value. The final bottom row corresponds to the joint distributions for q vs $\chi_{\rm eff}$, M vs $\chi_{\rm eff}$, and $\chi_{1,z}$ vs $\chi_{2,z}$ for all five injections.

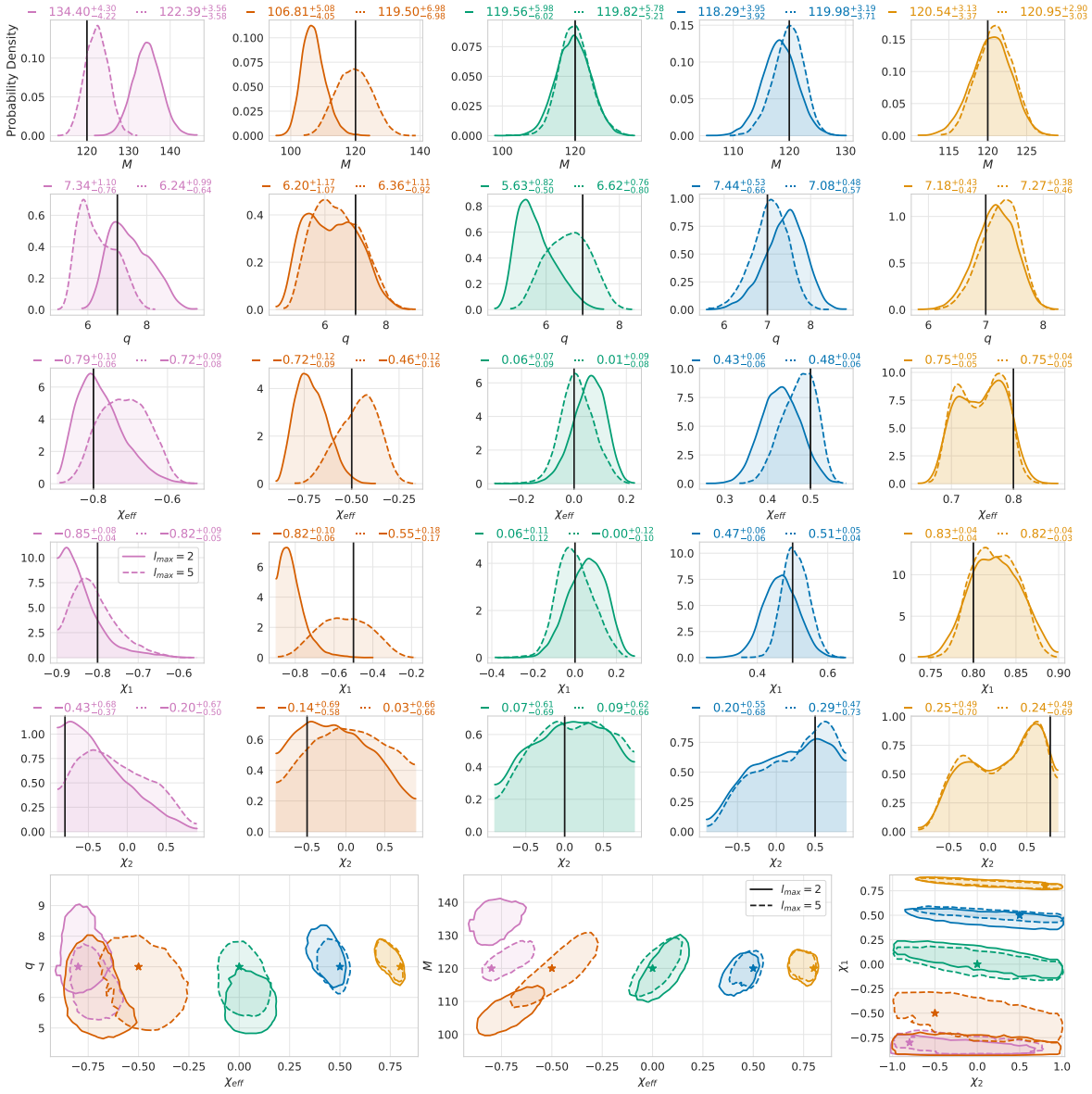

Figure 3.15: **Non-HM and HM runs for q=7 spin set, with SNR=30 and M=120**: The first five rows show the $M, q, \chi_{\rm eff}, \chi_{1z}, \chi_{2z}$ one-dimensional marginal distributions, where among this set of figures each column corresponds to a different synthetic source recovered with either all $\ell_{\rm max} = 5$ modes (dashed line) or $\ell_{\rm max} = 2$ modes (solid line). Our figures are organized such that the injected spin is systematically increased from left to right, where the synthetic source runs are ID12 ($\chi_{\rm eff} = -.8$), ID13 ($\chi_{\rm eff} = -.5$), ID14 ($\chi_{\rm eff} = 0$), ID15 ($\chi_{\rm eff} = .5$), and ID16 ($\chi_{\rm eff} = .8$). In each figure's title, we report the median value and the 90% confidence intervals of the marginalized 1D distribution for the $\ell_{\rm max} = 2$ (left) and $\ell_{\rm max} = 5$ (right) cases. A solid black vertical line denotes the true parameter value. The final bottom row corresponds to the joint distributions for q vs $\chi_{\rm eff}$, M vs $\chi_{\rm eff}$, and $\chi_{1,z}$ vs $\chi_{2,z}$ for all five injections.

and 3.17 show the posterior distributions for ID4 and ID8 respectively. As the SNR increases, the posteriors become more precise for both the non-HM and HM results (i.e., the statistical errors get smaller). However, the HM results converge on the true parameters while the non-HM results converge to a point offset from the true parameter (i.e., the systematic errors remain the same size and will dominate the statistical uncertainties). As GW detectors get more sensitive, the need for HM will become paramount even for the simplest of events. More sensitive detectors will potentially bring into view more exotic configurations at low SNRs which can also be problematic. For example, the weakest $q = 4$, $\chi_{\rm eff} = -0.5$ system has noticeable bias. This could be anticipated by noting that the mismatch between $\ell_{\rm max} = 5$ and $\ell_{\rm max} = 2$ models at this injection value is 0.06989 and so Eq. (3.4.6) is not satisfied.

One particularly challenging configuration was the loudest $q = 4$, $\chi_{\rm eff} = -0.5$ system shown in Figure 3.17 (solid blue). In particular, the posterior recovered with the $\ell_{\rm max} = 2$ model shows evidence for a secondary peak widely separated from the primary one. We checked this unexpected feature by directly comparing the values of the likelihood in a small neighborhood around both peaks. The presence of these two widely-separated peaks proved to be challenging for the current implementation of the ILE/RIFT algorithm, which uses a single interpolant of the log-likelihood surface. As a result, running this case took a significantly longer time while also achieving a comparatively lower accuracy, where the accuracy is quantified by the effective number of adaptive Monte Carlo samples. This case underscores that for high SNR events the omission of subdominant modes can introduce highly complex likelihood surfaces, and prove challenging to explore accurately. Within the RIFT framework, a recently implemented Gaussian Mixture Model sampler is expected to more efficiently sample from complicated likelihood surfaces. This case also demonstrates how incorrect models can accidentally yield good recovery of some parameters: the marginalized posterior for χ_2 (solid blue curve) looks remarkably accurate around the primary peak despite the joint posterior (bottom right panel) being nowhere near the true value.

To quantify the bias between the non-HM and HM runs, we consider two commonly used measures of bias: (i) classifying the recovery of a particular parameter as biased if the injected

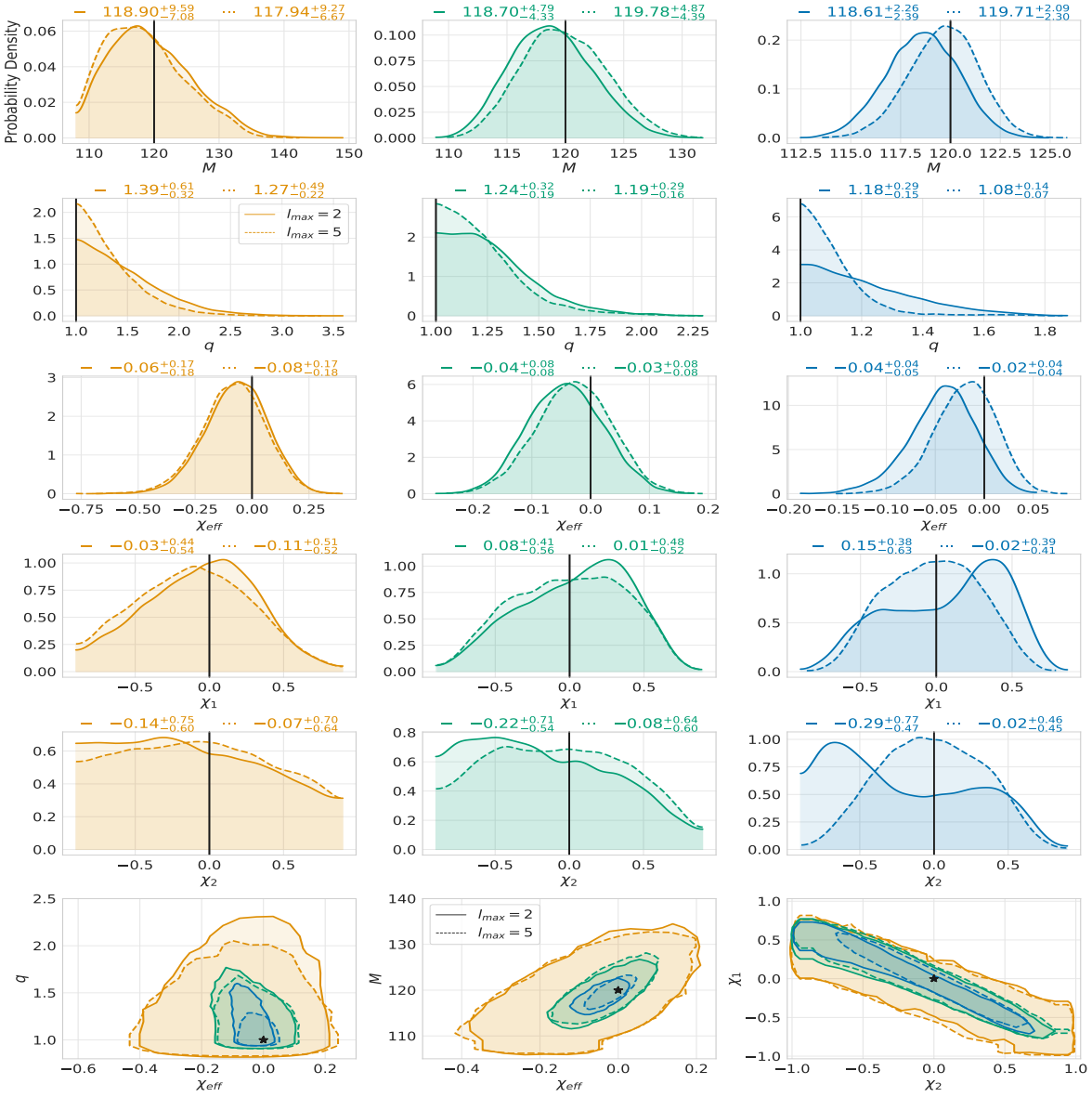

Figure 3.16: **Non-HM and HM runs for a q=1, M=120, and zero-spin source (ID4), for different SNRs**: The first five rows show the $M, q, \chi_{\rm eff}, \chi_{1z}, \chi_{2z}$ one-dimensional marginal distributions, where among this set of figures each column corresponds to a different synthetic source recovered with either all $\ell_{\rm max} = 5$ modes (dashed line) or $\ell_{\rm max} = 2$ modes (solid line). Our figures are organized such that the signal's network SNR is systematically varied as 10 (orange), 30 (green), and 70 (blue), corresponding to the left, middle, and right columns, respectively. A solid black vertical line denotes the true parameter value. The final bottom row corresponds to the joint distributions for q vs $\chi_{\rm eff}$, M vs $\chi_{\rm eff}$, and $\chi_{1,z}$ vs $\chi_{2,z}$ for all three injections.

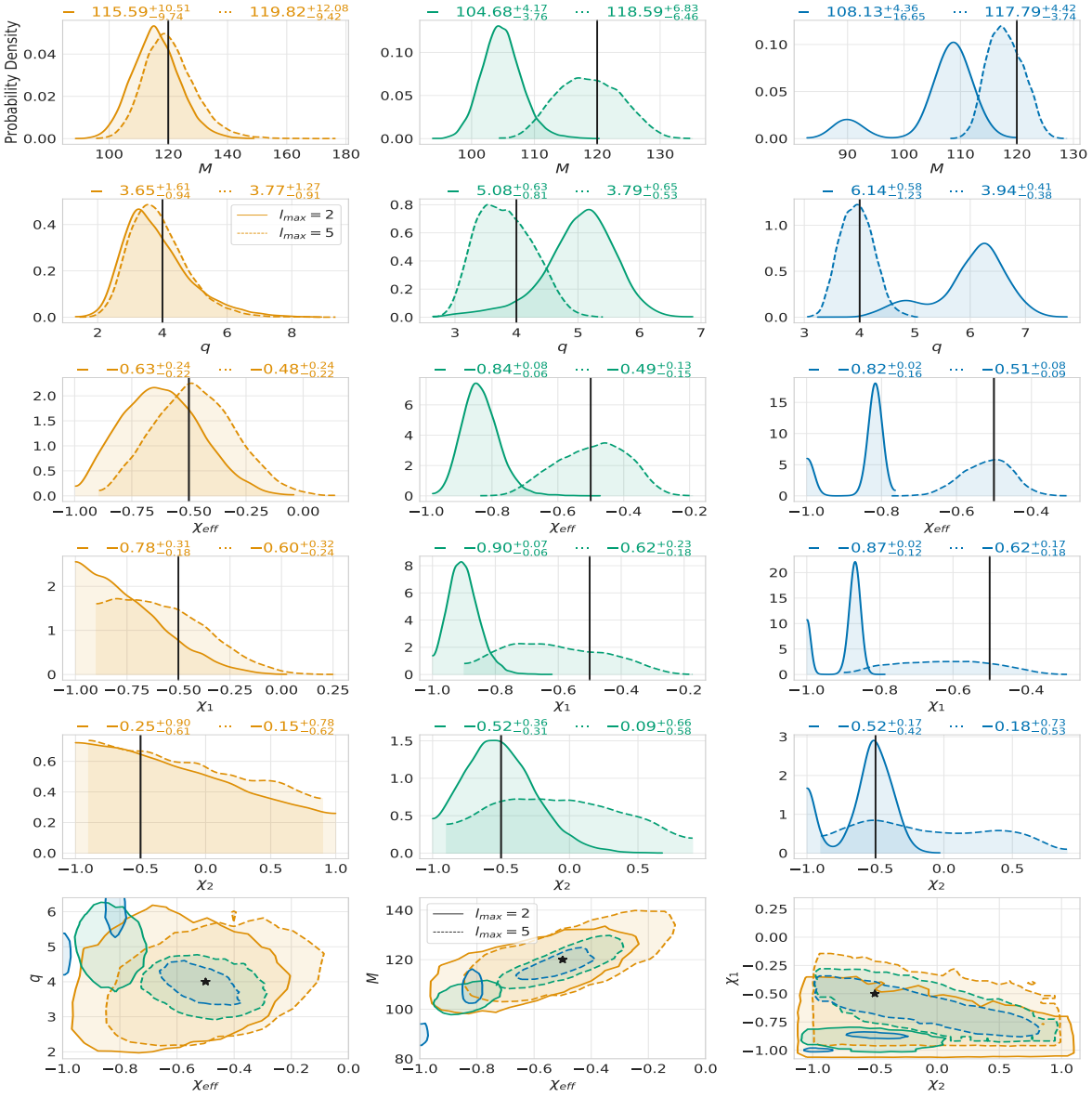

Figure 3.17: **Non-HM and HM runs for a q=4, M=120, and $\chi_{\rm eff}$ = -0.5 source (ID8), for different SNRs**: The first five rows show the $M, q, \chi_{\rm eff}, \chi_{1z}, \chi_{2z}$ one-dimensional marginal distributions, where among this set of figures each column corresponds to a different synthetic source recovered with either all $\ell_{\rm max} = 5$ modes (dashed line) or $\ell_{\rm max} = 2$ modes (solid line). Our figures are organized such that the signal's network SNR is systematically varied as 10 (orange), 30 (green), and 70 (blue), corresponding to the left, middle, and right columns, respectively. A solid black vertical line denotes the true parameter value. The final bottom row corresponds to the joint distributions for q vs $\chi_{\rm eff}$, M vs $\chi_{\rm eff}$, and $\chi_{1,z}$ vs $\chi_{2,z}$ for all three injections.

parameter value is outside of the 90% confidence region and (ii) the Jensen-Shannon divergence (JSD) between the different parameter distributions. Given two probability distributions $p(x)$ and $g(x)$, the JSD is defined as

$$D_{\rm JS}(p\,|\,g) = \frac{1}{2}\Big(D_{\rm KL}(p\,|\,s) + D_{\rm KL}(g\,|\,s)\Big)\,, \tag{3.4.7}$$

where $s = 1/2(p+g)$ and

$$D_{\rm KL}(p\,|\,g) = \int p(x)\log_2\left(\frac{p(x)}{g(x)}\right)\mathrm{d}x\,, \tag{3.4.8}$$

is the Kullback-Leibler divergence (KLD) between the distributions p and g, measured in bits. For context, this is the same calculation the LVC performed in [15] to quantify the agreement between different models. When measured in bits, the JSD is bounded below by 0. For a sense of scale, the KL divergence between two one-dimensional Gaussians with identical standard deviations but differing means μ_1, μ_2 is $(\mu_1-\mu_2)^2/2\sigma^2\ln 2$; inverting, JSD $= 0.2$ corresponds to $\mu_1 - \mu_2 \simeq 0.5\sigma$.

Figure 3.18 shows the JSD vs SNR and the simple “bias classifier" for both the ID4 and ID8 runs, respectively. Following the discussion in the LSC's recently published Gravitational-Wave Transient Catalog [15] (cf. Appendix 2.B), we consider two marginalized posteriors to be sufficiently different (i.e. biased) if the JSD is greater than ≈ 0.15. This number corresponds to a SNR $\simeq 30$ for non-spinning, equal-mass binaries; SNR $\simeq 10$ at $q = 4$ and $\chi_{1z} = \chi_{2z} = -0.5$. Since subdominant modes become more important at larger mass ratios and more negative values of $\chi_{\rm eff}$, the quoted SNRs provide convenient lower bounds for similar systems. For example, we expect HMs will also affect the posterior for systems with SNRs ≥ 30 and $q > 1$, $\chi_{\rm eff} \leq 0$ (similar to ID4); for systems with SNRs ≥ 10 and $q > 4$, $\chi_{\rm eff} \leq -0.5$ (similar to ID8).

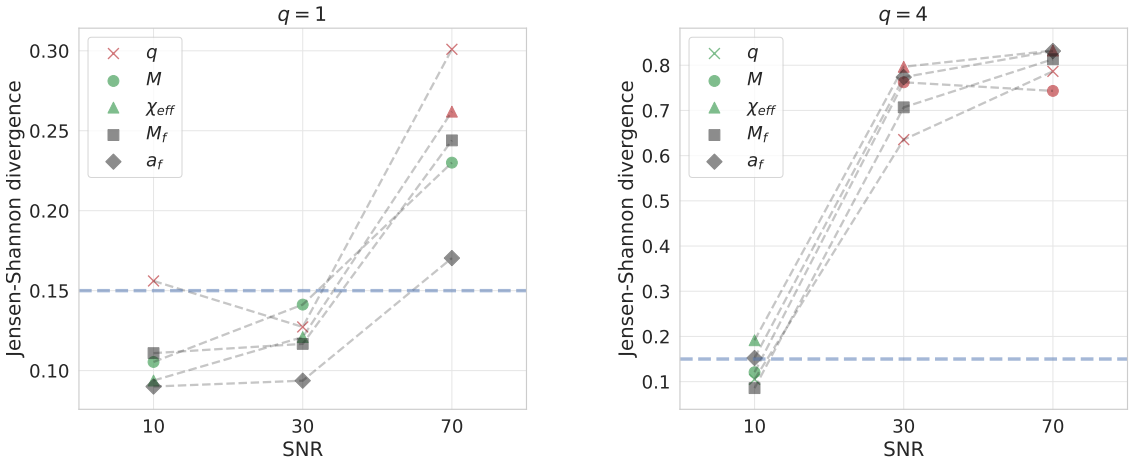

Figure 3.18: **The importance of higher modes for loud signals: bias vs SNR**: These panels show the JSD vs SNR for source ID4 (left panel) and ID8 (right panel). Different markers indicate which one-dimensional marginal distribution was used to evaluate the JSD, which are depicted in Figures (3.16) and (3.17) for ID4 and ID8, respectively. The dashed horizontal blue line demarcates a commonly used threshold for unacceptably large bias. Markers colored in red indicate that the true value falls outside the 90% credible interval region for the $\ell_{\rm max} = 2$ case (significant bias in the recovered parameter value), while those colored in green indicate the opposite. For $\ell_{\rm max} = 5$, the true value is almost always within the 90% credible interval region except the parameter q in the $q = 1$ case, where the true value lies at the edge; despite not being in the he 90% credible interval the marginalized distribution for q obtains its maximum value at $q = 1$ (cf. row 2 of Figure 3.16). Markers in gray indicate the JSD for the final remnant masses and spins.

3.4.4 Discussion

3.4.4.1 Effect of Different Spin Priors

Besides the impact of sub-dominant modes, the ability to accurately measure the spin parameters is also influenced by the choice in spin prior [170], which is not well-informed by astrophysical observations or source population models. In our study, we have used a prior which is uniform in $\chi_{\rm z}$ (P1). However, many of the LVC's analysis assume a prior that is uniform in spin magnitude, $|\vec{\chi}|$, and on the 2-sphere, which, for our non-precessing model, would induce a prior by projection of $\vec{\chi}$ along the orbital angular momentum vector (P2). When assuming this spin prior, the peak of the PDF of the individual component spins will strongly favor zero. To see how these two significantly different priors affect the ability to measure the spins, we compare posteriors for two runs ID2 and ID6 with SNR= 30 assuming the two different priors. Figure 3.19 shows the individual $\chi_{*\rm z}$ spins as well as the effective spin $\chi_{\rm eff}$ for each spin prior. Despite using a strong source, all spin parameters are significantly perturbed by the prior choice, similar to results found in previous work [170].

3.4.4.2 Consequences of biases for remnant properties and consistency tests

Using the posterior distributions of the BBH system's component masses and spins one can compute the remnant mass, M_f, and spin, a_f of the final (merged) black hole. The values of (M_f, a_f) are interesting in their own right as they can be used to infer a population of astrophysical compact objects that formed through the merger of a BBH system. Another use of (M_f, a_f) is to test the consistency of general relativity by predicting these remnant values found from (i) the post-merger portion of the signal which is described by a ringdown signal characterized entirely by (M_f, a_f) and (ii) the inspiral portion of the signal where we compute the BBH system's component values and, using numerical relativity, arrive at an alternative estimate of (M_f, a_f). If general relativity correctly describes the system's entire evolution, we should expect the remnant values found through each to be mutually consistent [171]. A closely related test uses the remnant values computed with the inspiral-only portion of the signal to infer the expected quasi-normal mode (QNM) of ringdown signal, and then comparing this

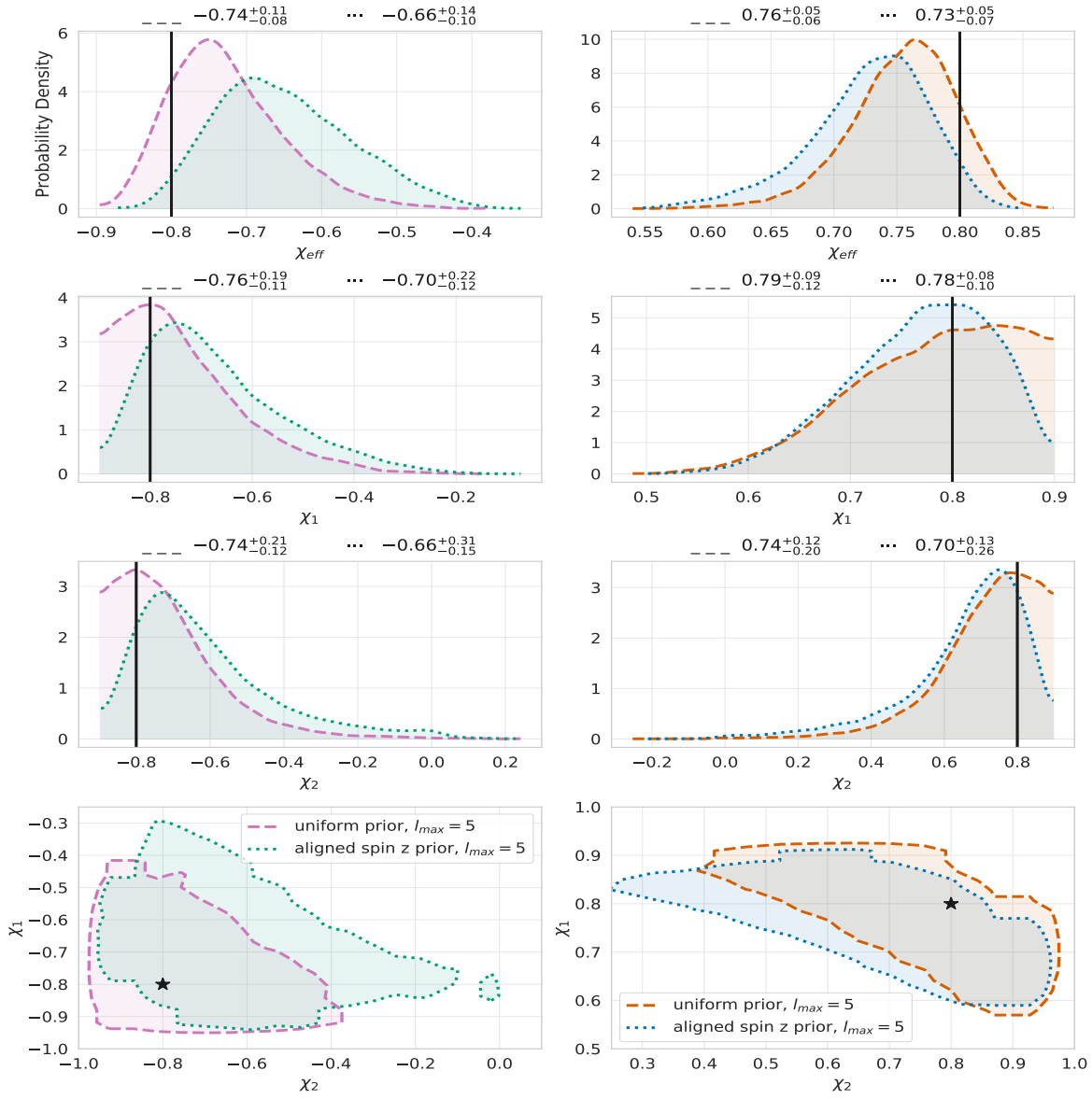

Figure 3.19: **The effect of priors on spin measurability**: Individual and effective spin parameter recovery assuming two different priors, using synthetic datasets ID2 ($q = 1$, $\chi_{\rm eff} = -.8$) and ID6 ($q = 1$, $\chi_{\rm eff} = .8$) with SNR= 30. The dashed curve represents the results using a prior that assumes uniform spin magnitudes in χ_z (P1; uniform prior), and the dotted curve represents the results using a prior that assumes uniform spin magnitudes in $\vec{\chi}$ (P2; aligned spin z prior). Despite the high value of SNR used here, we observe that the choice of prior has a significant influence on the recovered posteriors.

predicted QNM spectrum with the QNMs estimated directly from the ringdown-only portion of the data [171]. A different, but related, set of tests of the no-hair theorem also benefit from the inclusion of both higher harmonics and as well as quasinormal mode overtones [172].

All of these studies require accurate measurement of the system's remnant masses and spins. For example, unacceptably large bias in these quantities could provide misleading evidence for failed GR consistency tests, unless the quadrupole-only pre-merger and post-merger models make a serendipitously incorrect inference of the remnant properties (i.e., both models are incorrect but in a consistent manner).

In this subsection we explore bias in the remnant properties implied by the posterior distributions computed in Section 3.4.3 as the SNR increases. We compute the remnant mass and spin magnitude by evaluating the high-accuracy fitting formula provided by the surfinBH Python package [173] on the posteriors computed using $\ell_{\rm max} = 5$ and $\ell_{\rm max} = 2$ recovery waveform models.

As the first example, where we expect minimal bias, we consider the $q = 1$, zero-spin source system (ID4) whose posterior distributions for SNRs= $\{10, 30, 70\}$ are reported in Figure 3.16 from which we compute remnant posteriors in Figure 3.20 (left set of figures). While the true remnant values are contained within all of the joint posteriors's 90% credible region, we begin to see modest bias indicating impact from the higher-modes when the signal's strength reaches an SNR value of 70. This is quantified in Figure 3.18 which shows the Jensen-Shannon divergence for M_f and a_f are 0.24 and 0.17, respectively. For context, values above 0.15 are typically considered to reflect non-negligible bias [15]. At all values of the SNR, we find the $\ell_{\rm max} = 5$ posterior more tightly constrains the true values.

Figure 3.20 also shows a similar sequence for the $q = 4$, $\chi_{\rm eff}$ = -0.5 source (ID8) where now the true remnant values are no longer contained within the 90% credible intervals by SNR=30. As seen from Figure 3.18, the JS divergence is already close to, or greater than, 0.15 at SNR=10. This suggests that higher modes are very important when estimating the remnant values from such systems, and neglecting them would incorrectly lead to a failure of the IMR consistency test for essentially any event we might conceivably observe similar to

ID8.

3.4.4.3 Consequences of biases on population reconstruction

In a second and more qualitative example of the impact of parameter biases due to neglect of physics, we consider astrophysical inference for the mass, mass ratio and spin distribution of coalescing BHs. For example, consider an SNR=30, zero-spin BBH event with $q = 4$. As illustrated by the green curves in Figure 3.14, inferences which neglect HMs would deduce negative effective spin (and a more extreme mass ratio). A single source with definitively negative $\chi_{\rm eff}$ would be interpreted as a strong indication for dynamical formation in samples of less than several hundred mergers. Such biased inferences for high-amplitude sources could thus be misinterpreted to support qualitatively different formation channels (e.g., dynamical formation) than supported by the true parameters, which are well-characterized by multimodal PE.

More typically, parameter biases due to model incompleteness enter more insidiously into astrophysical inference, since population inference relies on combining information from multiple sources and since systematic biases impact all sources at a similar level. Following [111], we estimate that parameter biases $\Delta x = x^{\rm true} - x^{\rm median}$ will be significant for a population of N sources if the bias can be identified in the population mean by stacking observations: in other words, if $\Delta x \gtrsim \sqrt{\sigma^2_{stat} + \sigma^2_{astro}}/\sqrt{N}$ where σ_{stat} and σ_{astro} are the statistical error in x and the width of the astrophysical distribution of x, respectively. In terms of the JSD we anticipate that systematic differences in waveforms must produce a change in posteriors less than JSD $= 0.15/N$ to have no effect on population inference. Our examples show that even for zero-spin (but unequal-mass) binaries, inferences about the mass ratio, total mass, and effective spin in moderate-SNR sources can be significantly biased by the lack of HOMs. If a population of unequal mass-ratio binaries exist and has a spin distribution qualitatively similar to the seemingly low-spin BH population identified in O2, even inferences drawn from a handful of observations could be noticeably biased about BH masses and spins.

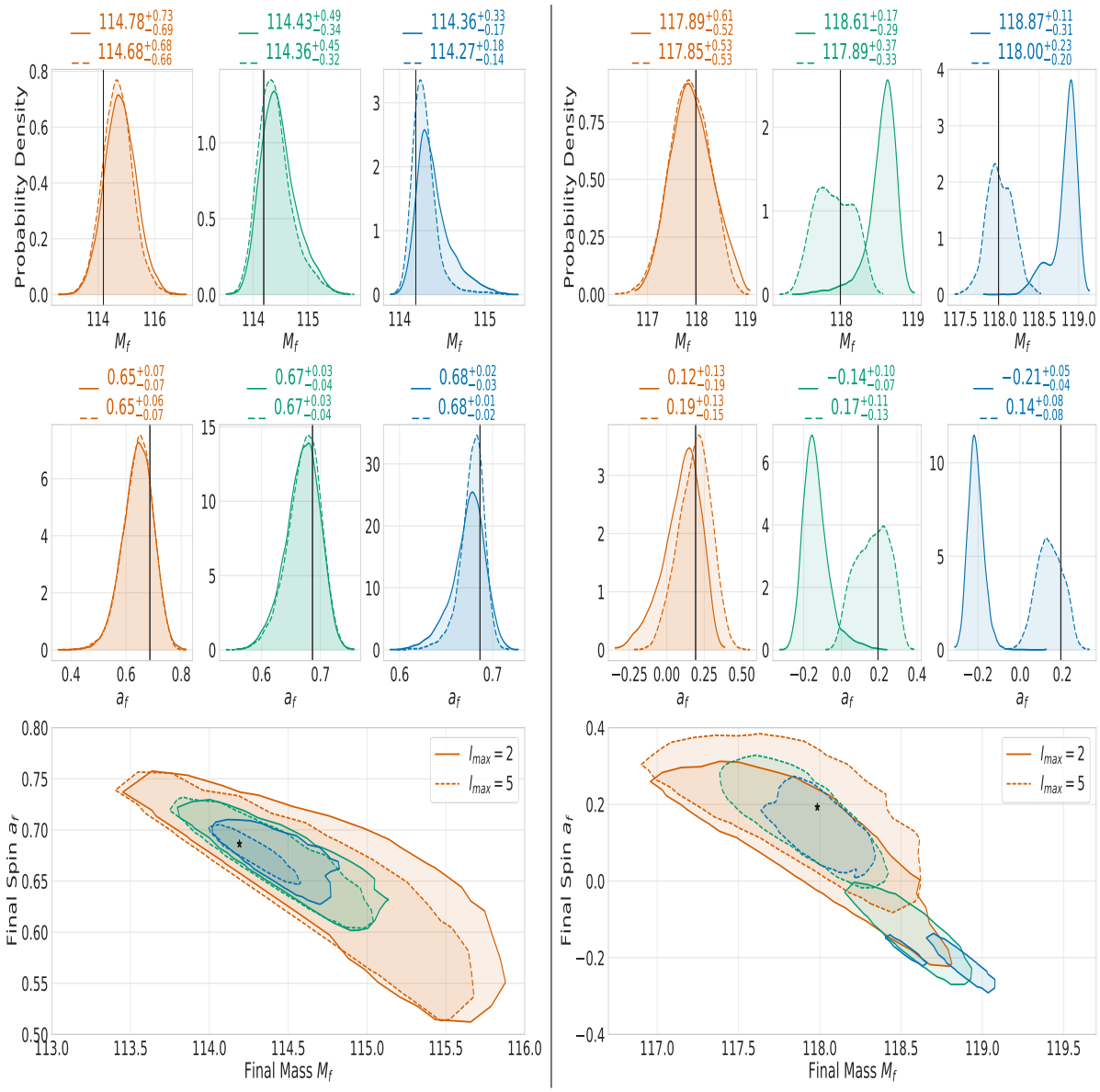

Figure 3.20: **Effect of higher-order modes on remnant values and IMR consistency tests**: These panels show marginal distributions for remnant properties of the redshifted mass, M_f, and spin, a_f, for a non-spinning, $q = 1$ source (ID4; left panels) and $\chi_{\rm eff} = -0.5, q = 4$ source (ID8; right panels). Our figures are organized such that the signal's network SNR is systematically varied as 10 (orange), 30 (green), and 70 (blue), corresponding to the left, middle, and right columns of each panel.

3.4.4.4 A GW170729-like source

While much of our focus has been on fiducial BBH systems, it is also interesting to consider sources that are similar to events from the most recent observing run. In this subsection, we analyze a synthetic source that has parameters (cf. ID1 in Table 3.6) similar to GW170729, one of the more interesting events from O2. As mentioned in [15] and [167], the SNR of GW170729 was ~ 12. However, to better highlight the importance of HMs for this event, we instead consider a GW170729-like event located at a distance such that the SNR is 30. For consistency with other synthetic events analyzed throughout this paper, we set $\chi_{\rm eff} = 0.5$ as its true value, which is near the upper end of the 90% credible interval t [174]. Note that although $\chi_{\rm eff} = 0.5$, we now have $\chi_{1z} \neq \chi_{2z} = 0$. We continue using a uniform spin magnitude in $\chi_{\rm z}$ as our spin prior.

Figure 3.21 shows the posterior distributions for the runs that include only $\ell_{\rm max} = 2$ (solid lines) and include all the $\ell_{\rm max} = 5$ (dashed lines). As with all the results in Section 3.4.3, we see a significant bias between the two runs in all the parameters. For example, we see that the $\ell_{\rm max} = 5$ model does a much better job at recovering the individual spin components as well as placing somewhat tighter constraints on the spin of the larger BH, χ_{1z}. Interestingly, we see a similar shift in q and $\chi_{\rm eff}$ that was observed in a recent re-analysis of the actual GW170729 event [174]. As our detectors continue to get more sensitive, we will increasingly see events with parameter and SNR values similar to the synthetic source ID1 considered here.

3.4.4.5 Measuring individual black hole spins

It is well known that while individual spins are difficult to measure, the effective spin parameter, $\chi_{\rm eff}$, is much better constrained. A recent study [175] systematically explored this question in the context of a single gravitational-wave detector by using the quadrapole-only SEOBNRv2 model [54, 39]. The general conclusion of this work (see Figures 1 and 4 of Ref. [175]) is that individual spins are poorly constrained. For equal-mass systems, it was found that the spin measurements are constrained only by the Kerr limit and so only near-extremal spins can be constrained as the posterior will run up against the prior. Furthermore, as the mass ratio

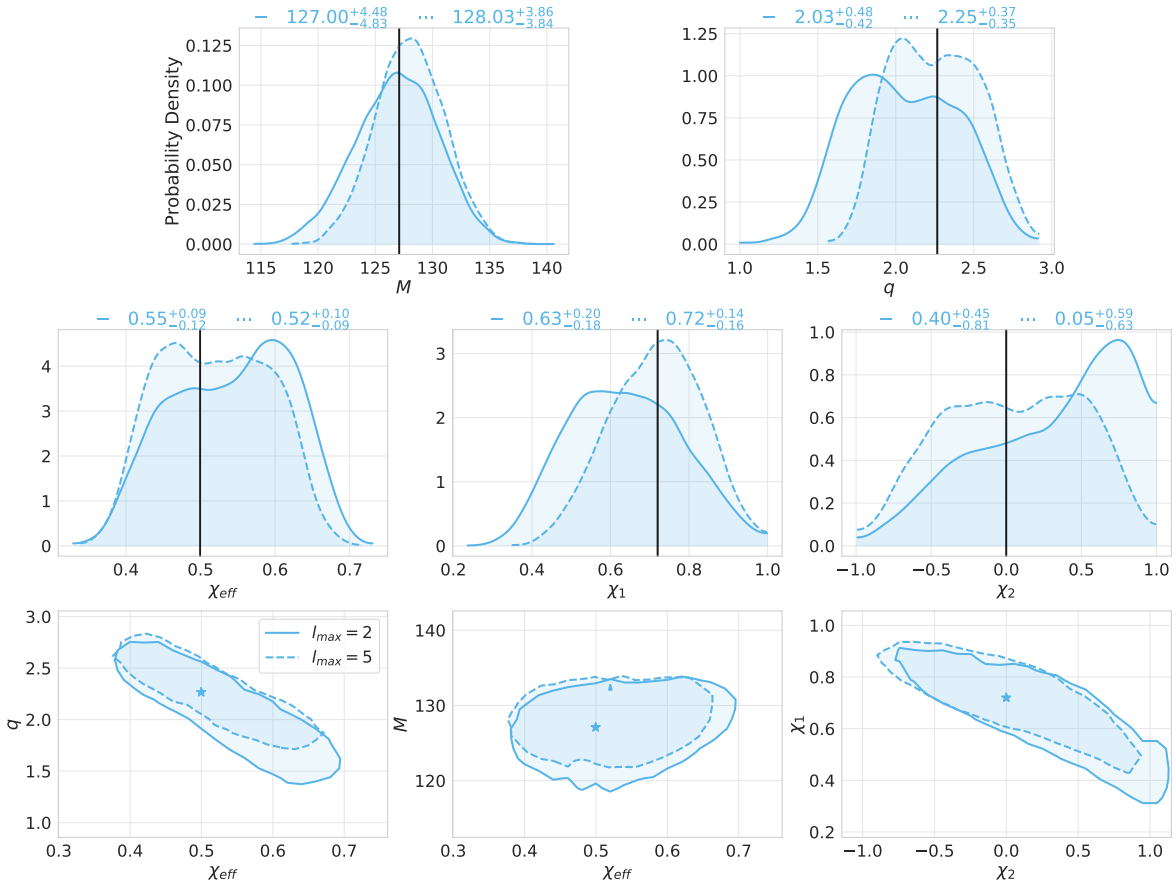

Figure 3.21: **GW170729-like event**: Posterior plots for the ID1 run: $q = 2.267$, $M(M_\odot) = 127.1$, $\chi_{1z} = 0.72$, $\chi_{2z} = 0.0$, SNR= 30. The solid and dashed lines represent the $\ell_{\rm max} = 2$ and $\ell_{\rm max} = 5$ runs respectively. When including HM, we are able to improve the recovery of individual spin components. We also see a significant shift in the q and $\chi_{\rm eff}$ distributions.

increases, the spin of the larger blackhole is better constrained while the smaller black hole's spin remains unconstrained. Finally, this general picture remains unchanged across a wide range of total masses, including the values we have focused on in our paper.

In this subsection, we revisit the results from Section 3.4.3 but now briefly comment on our ability to measure the individual component spins using the full three-detector network with a our multi-mode recovery model.

Unfortunately, as anticipated in Ref. [175], the inclusion of subdominant modes does not qualitatively change the situation. This is visually and quantitatively evident for equal mass (cf. Figure 3.13), $q = 4$ (cf. Figure 3.14), and $q = 7$ (cf. Figure 3.15) systems, all of which have a network SNR of 30. Here we see that while the inclusion of subdominant modes (dashed lines) dramatically reduces the bias in recovering $\chi_{\rm eff}$, χ_1, and χ_2, the size of the 90% confidence intervals (shown in the figure's title) are mostly unaffected. A similar conclusion can be reached by comparing the joint distributions for χ_1 vs χ_2 (bottom right panels in Figures 3.13, 3.14, and 3.15) recovered with $\ell_{\rm max} = 2$ and $\ell_{\rm max} = 5$ recovery models.

Thus we conclude that, at least for the configurations considered here, including subdominant modes in our waveform recovery model will reduce bias in the both the effective spin and individual spin components, but does relatively little to better constrain them.

3.4.5 Appendix: Follow up on the significance of higher modes for equal mass, zero spin, SNR$= 10$ case

As pointed out in Sections 3.4.3.4, there seems to be significant differences between the $\ell_{\rm max} = 5$ and $\ell_{\rm max} = 2$ runs for the equal mass, zero spin, SNR$= 10$ case, which runs contrary to several previous studies that had implied that HM would have minimal impact at low SNR for comparable-mass binaries.

To better understand our results, we perform a complementary analysis under the assumption of zero spin (i.e. lay out a grid only in $M_{\rm tot}, q$), allowing us to directly evaluate the marginal likelihood versus the two remaining binary parameters. Figure 3.22 shows the results of both the $\ell_{\rm max} = 5$ and $\ell_{\rm max} = 2$ results. We continue to observe notable differences between

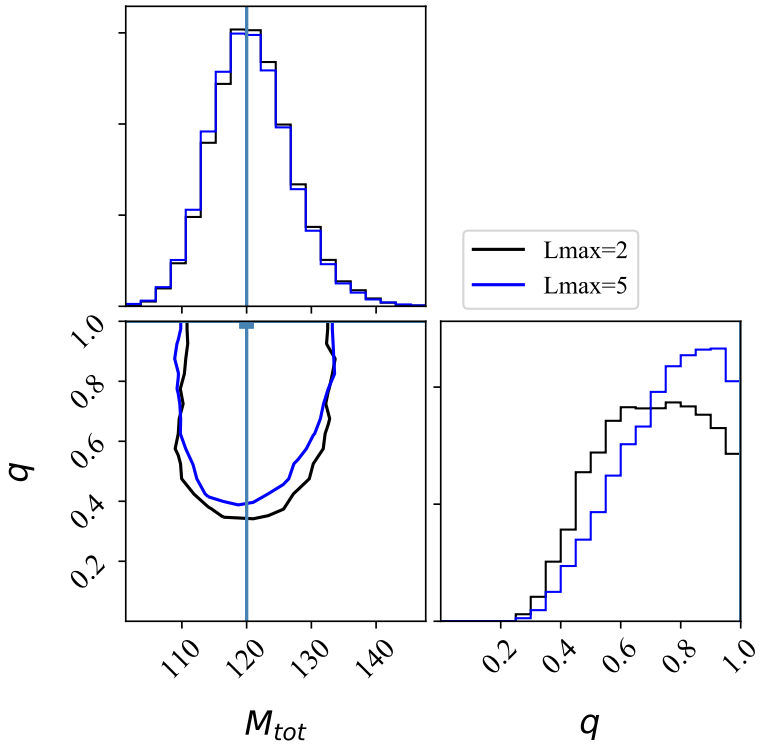

Figure 3.22: **Reanalysis of equal mass, zero spin, SNR**= 10: This corner plot shows the reanalyses of a equal mass, zero spin, SNR= 10 source using $\ell_{\rm max} = 2$ (black) and $\ell_{\rm max} = 5$ (blue) mode but only on a grid in mass parameters (i.e. assuming zero spin). As first shown in Figure 3.16, there are noticeable differences between the two different distributions.

the two posteriors even when restricted to two dimensions (i.e. only mass parameters). It is certainly surprising to see any difference given that this is a low SNR, equal mass event. One possibility is that due to the broadness of the posterior in mass ratio, a significant fraction of the posterior needs to be evaluated at values of $q \gtrsim 2$ where higher modes begin to play an increasingly important role.

3.5 Inferring parameters of the loudest intermediate mass black hole trigger in LIGO's O1/O2 data

In this work (see also related paper [5]), we reanalyzed the loudest intermediate mass binary black hole trigger from LVC's O1/O2 data. In particular we again use RIFT concurrently NR simulation-based waveform models that includes radiated higher order modes $(\ell, |m|)$ beyond the dominant $(2, 2)$ mode [57, 176]. These higher order modes are critical for analysis of such massive binary black hole mergers [177]. In particular, we use the two waveform models: (i) `NRHybSur3dq8`, which is our preferred model in this study [57], and (ii) `NRSur7dq4`, which we utilized for additional checks (i.e. information about transverse spins) [176]. While both models are tuned directly to NR simulations, Model (i) includes hybridized waveforms that are valid up to a mass ratio $m_2/m_1 \geq 1/8$ and for binary black holes with aligned spins $|\chi_{1z,2z}| \leq 0.8$ (in dimensionless units) but extrapolates well to $|\chi_{1z,2z}| \leq 0.9$ while Model (ii) does not include hybridized waveforms but is valid up to mass-ratio $m_2/m_1 \geq 1/4$ and for binary black holes with generic spin orientation that captures the spin-orbit precession with the same spin magnitude restrictions. In context of GW170502, the specific advantage of using these two models is the inclusion of the radiated higher order modes up to $\ell \leq 4$ and $(\ell, |m|) = (5, 5)$ (the latter only for Model (i)).

Along with GW170817A [16] and GW151205 [17], the candidate GW170502 adds to an emerging population of high mass binary analyzed in the open data era. All of these open-data events have similar statistical significance, i.e., they will be considered *marginal* candidates in comparison to the confirmed LIGO-Virgo events. However, no such triggers in gravitational wave astronomy are unique, meaning if we see it once, we will see something similar again;

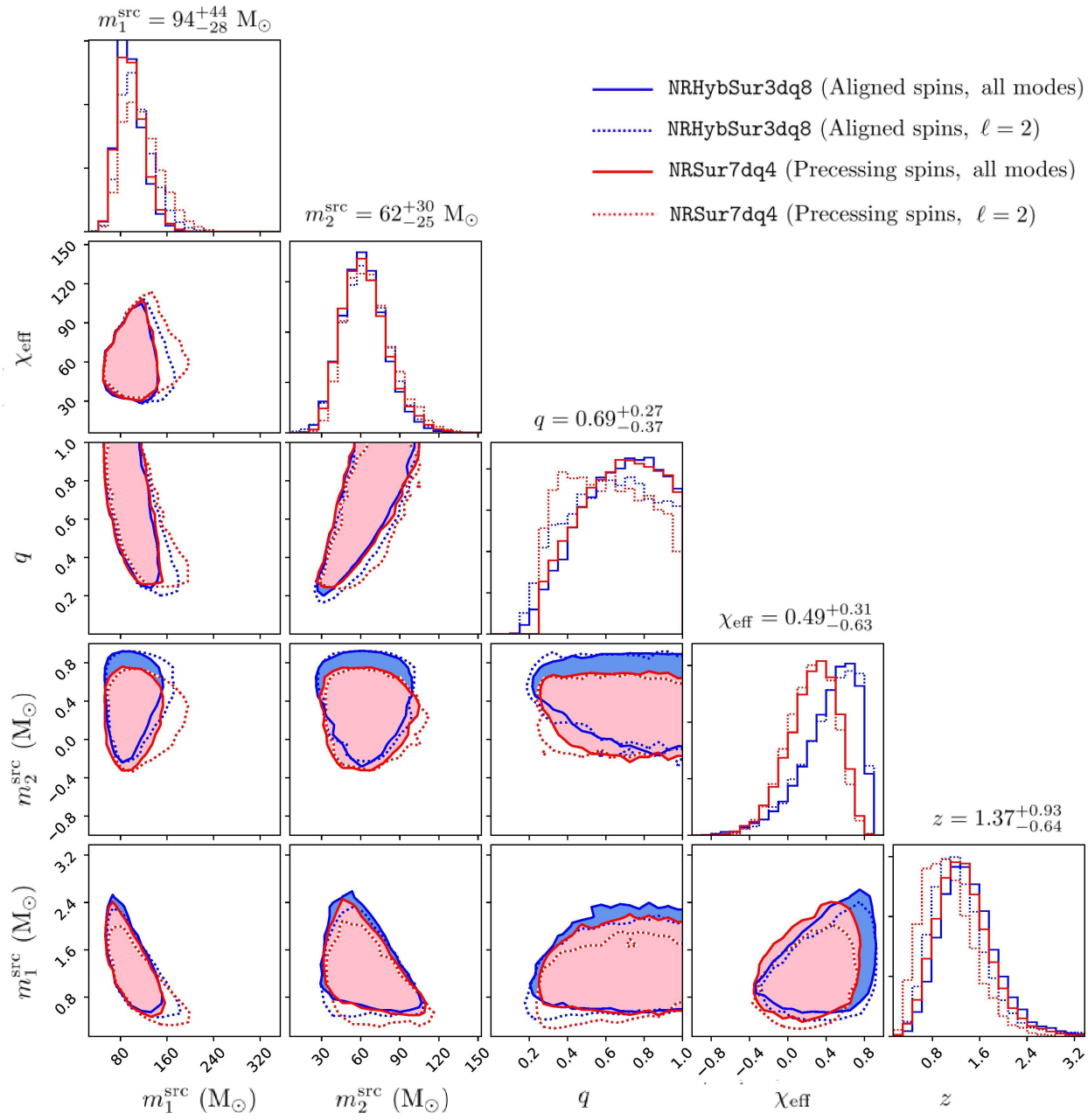

Figure 3.23: Marginalized posteriors for GW170502 using `RIFT`. Two-dimensional contours enclose 90% of the distribution. The two colors refer the two waveform models `NRHybSur3dq8` (blue) and `NRSur7dq4` (red). The solid lines refer to results for including all the higher order modes in the waveform (`NRHybSur3dq8` with $\ell \leq 5$ and `NRSur7dq4` with $\ell \leq 4$), the dotted line restrict models to $\ell = 2$. The numbers quoted above each column are the median, with the 90% interval obtained from `NRHybSur3dq8` using all the available higher order modes.

Waveform Model	`NRHybSur3dq8` (aligned spins)		`NRSur7dq4` (precessing spins)	
Radiated Modes (ℓ, m)	$\ell \leq 4$, $(5,5)$	$\ell = 2$	$\ell \leq 4$	$\ell = 2$
Primary BH mass, $m_1^{\rm src}$ [$M_\odot$]	94^{+44}_{-28}	104^{+56}_{-35}	96^{+45}_{-30}	112^{+66}_{-40}
Secondary BH mass, $m_2^{\rm src}$ [$M_\odot$]	62^{+30}_{-25}	63^{+32}_{-27}	62^{+34}_{-24}	65^{+36}_{-24}
Total mass, $M_{\rm tot}^{\rm src}$	157^{+55}_{-41}	169^{+60}_{-48}	159^{+61}_{-44}	179^{+76}_{-53}
Chirp mass, $\mathcal{M}_c^{\rm src}$	65^{+24}_{-17}	69^{+25}_{-20}	66^{+27}_{-18}	73^{+30}_{-21}
Mass-ratio, $q = m_2/m_1$	$0.69^{+0.27}_{-0.37}$	$0.64^{+0.32}_{-0.36}$	$0.68^{+0.28}_{-0.35}$	$0.60^{+0.34}_{-0.31}$
Effective inspiral spin, $\chi_{\rm eff}$	$0.49^{+0.31}_{-0.63}$	$0.49^{+0.32}_{-0.62}$	$0.28^{+0.35}_{-0.49}$	$0.26^{+0.35}_{-0.45}$
Effective precession spin, χ_p	-	-	$0.60^{+0.24}_{-0.32}$	$0.62^{+0.23}_{-0.32}$
Redshift, z	$1.37^{+0.93}_{-0.64}$	$1.20^{+0.93}_{-0.59}$	$1.29^{+0.87}_{-0.64}$	$1.02^{+0.84}_{-0.57}$

Table 3.7: Parameters of GW170502 for the two waveform models and different combinations of modes discussed in this study.

and, therefore the purpose of this study is not to make new statements on the detection confidence of GW170502 itself, the event rate limit of candidates like it, or its particular implications for the black hole population. Further, the probability of astrophysical origin for such triggers is subject to change with better sensitivity of LIGO/Virgo detectors. Therefore, our goal is to demonstrate the power of a new machinery that allows us to infer the properties of intermediate mass black hole candidates ($\gtrsim 100\ M_\odot$) in the current generation of gravitational-wave detectors.

To test the impact of higher order modes, we conduct a separate estimation only including the dominant $\ell = 2$ modes as well as runs that including all higher order modes with both `NRHybSur3dq8` and `NRSur7dq4` models (see Figure 3.23 and Table 3.7). We adopt the conventional mass and distance priors for our Advanced LIGO/Virgo data analysis: a uniform mass density in the dtector frame and uniform un the cube of the luminosity distance with a maximum distance of 30000 Mpc. For our aligned spin analyses, we adopt a uniform prior for $\chi_{i,z} \in [-0.9, 0.9]$, where $\chi_{i,z}$ is the component of the black hole spins aligned with the angular momentum, and assume there is no in-plane spins components. For the precessing analyses, we adopt a spin prior where the spin vectors are uniformly distributed within the unit sphere for $\chi_i \in [-0.9, 0.9]$.

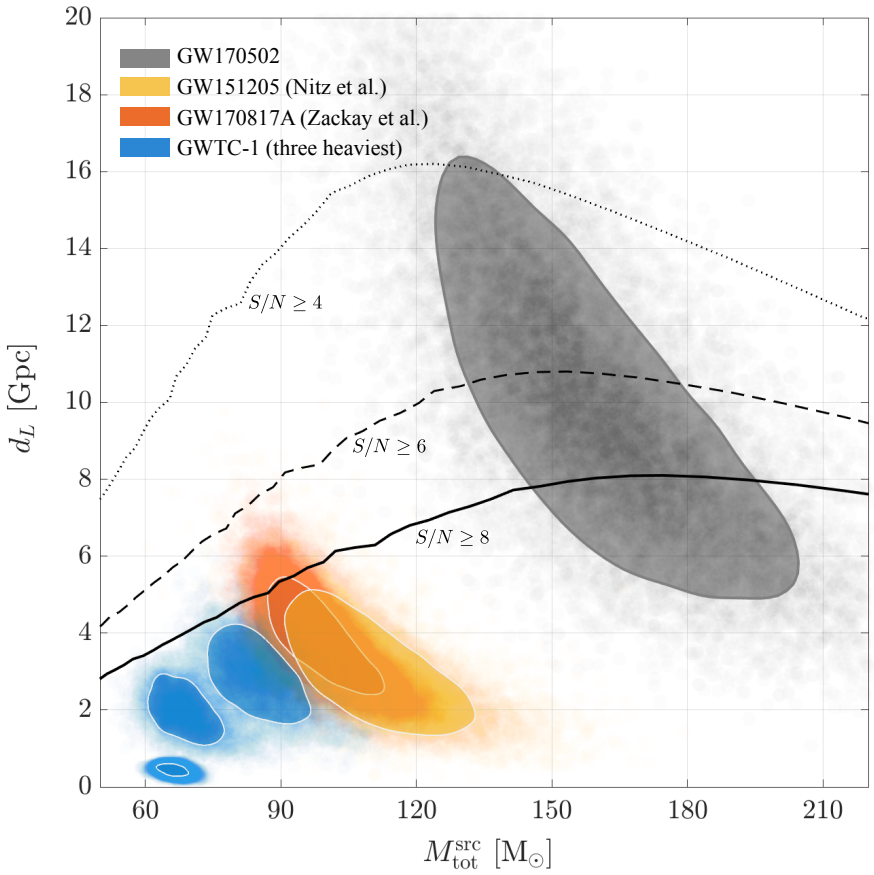

Figure 3.24: The six heaviest binary black hole mergers reported so far from the O1/O2 runs of Advanced LIGO/Virgo (2015-2017). The horizontal axis is the total mass in the astrophysical source frame and vertical axis is their corresponding luminosity distance. The contours refer to 90% confidence intervals and the transparent dots show the spread of the posterior sample. With the the black contour, we show the constraints on GW170502 using the `NRHybSur3dq8` model with all the available higher order modes. The blue contours show the three heaviest confirmed binary black hole mergers – GW170729, GW170823 and GW150914 – as reported in GWTC-1 [15]. In the orange and yellow contours, we show candidate GW170817A and GW151205 found by independent teams [16, 17]. The horizon distances for non-spinning, equal-mass binary black holes (black curves) are computed at different S/N for a single detector Advanced LIGO sensitivity during O2.

3.5.1 Results & Discussion

Masses By using all the available HOMs in the `NRHybSur3dq8` waveform model, we find that GW170502 corresponds to a total binary mass in the source frame of $M_{\rm tot}^{\rm src} = 157^{+55}_{-41}$ M$_\odot$ with 90% confidence interval. This makes the trigger heavier than all previous sources in the first

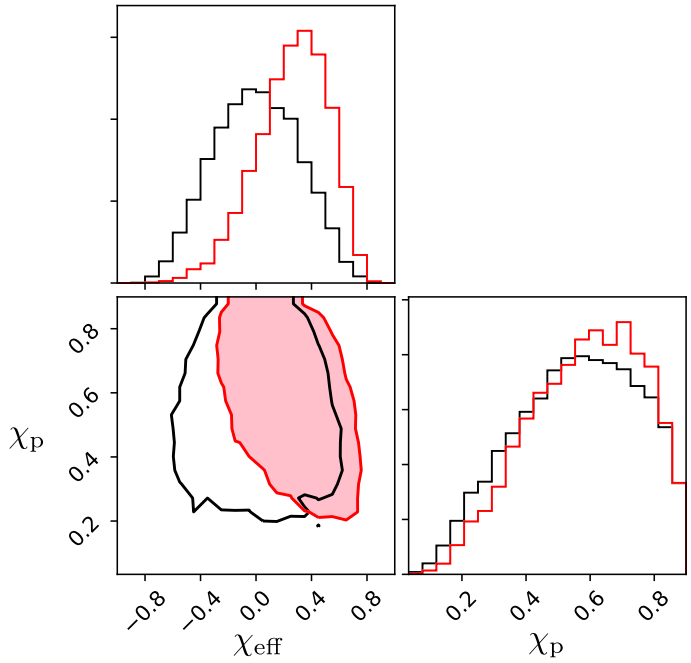

Figure 3.25: Marginalized posteriors of the effective inspiral spin parameter χeff and effective precession spin parameter χ_p) for GW170502 using `RIFT`. Two-dimensional contours show 90% intervals for `NRSur7dq4` model (red line) and the prior distribution (black line).

and second observing runs of Advanced LIGO (see Figure 3.24). The corresponding redshift and luminosity distance is constrained to $z = 1.37^{+0.93}_{-0.64}$ and $d_L = 10^{+8.9}_{-5.4}$ Gpc, which makes it potentially the farthest GW source. The observed trigger, therefore, was strongly redshifted, and the detector frame mass was about ~ 3 times heavier. The primary BH mass in the source frame was constrained to $m_1^{\rm src} = 94^{+44}_{-29}$ $\mathrm{M}_\odot$ and the secondary BH to $m_2^{\rm src} = 62^{+30}_{-25}$ $\mathrm{M}_\odot$. While the constraints on mass ratio are usually not as stringent with such massive binaries, we find that both waveform models put GW170502 at $q = m_2/m_1 \gtrsim 1/4$ within 90% confidence.

At the total mass of GW170502, the sensitivity of Advanced LIGO detectors in O2 at the detection threshold ($S/N = 8$) was up to ∼8 Gpc. Considering this trigger was sub-threshold with a network $S/N \sim 6$, it is reasonable that LIGO Livingston (the more sensitive detector) must have recorded $S/N \gtrsim 6/\sqrt{2} \sim 4$. This suggests GW170502 was well within the horizon volume (see Figure 3.24).

Waveform Model	NRHybSur3dq8		NRSur7dq4	
Radiated Modes	$\ell \leq 4$, $(5,5)$	$\ell = 2$	$\ell \leq 4$	$\ell = 2$
$\chi_{\rm eff} < -0.1$	6.0%	5.9%	9.4%	9.3%
$-0.1 < \chi_{\rm eff} < 0.1$	7.4%	7.8%	15.9%	18.3%
$0.1 < \chi_{\rm eff}$	86.6%	86.3%	74.7%	72.4%

Table 3.8: Probability from the posterior of each model and mode combination for value of $\chi_{\rm eff}$ in the specified bounds.

Spins For such a heavy binary, only the merger is essentially recorded in the LIGO frequency band. Therefore, the individual BH spins and their evolution remain ill-constrained. We, therefore, focus on constraining the effective inspiral spins, $\chi_{\rm eff}$, the net component of mass-weighted spins projected on the orbital angular momentum axis (for definition, see [26]). For GW170502, we constrain $\chi_{\rm eff} = 0.49^{+0.31}_{-0.63}$ with 90% confidence. The median and upper-bounds of $\chi_{\rm eff}$ are higher than reported earlier for this trigger (see Appendix-D of [178]). It is also significantly higher than all the BBH mergers of GWTC-1 (see Table III of [15]). For our preferred model `NRHybSur3dq8`, we find that the Bayes' Factor ($\mathcal{B}$) has a mild preference for spinning BHs over non-spinning ($\log_{10}\mathcal{B} = 0.46$).

Furthermore, Table 3.8 shows that in our preferred model, 86.6% of the posterior lies within the region $0.1 < \chi_{\rm eff}$. When taken with the fraction that lies below $\chi_{\rm eff} = -0.1$, we find that over 90% of the posterior lies outside the region $-0.1 < \chi_{\rm eff} < 0.1$, thus adding a strong support for a conclusion of non-zero effective inspiral spin. To investigate if the BH spins of GW170502 have components in the orbital plane, we measured the effective precession spin parameter $\chi_{\rm p}$ [179]. In Figure 3.25, we compare the constraints on the effective precession and inspiral parameter (χ_p vs $\chi_{\rm eff}$) for the fully precessing spin model `NRSur7dq4`. Similar to the aligned-spin model, we find $\chi_{\rm eff}$ consistently peaks at a positive value even if the model allows generic spin orientations. We calculate the Bayes' Factor $\mathcal{B}$ between precessing and aligned spin assumptions to be $\log_{10}\mathcal{B} = -0.68$ (disfavoring precession moderately) for the `NRSur7dq4` model, see Table 3.9. However, we gain no new information about χ_p and the spin-orbit precession of GW170502.

Model	NRHybSur3dq8	NRSur7dq4
$\log_{10}\mathcal{B}$ (HOM)	-0.03	0.00
$\log_{10}\mathcal{B}$ (spinning)	0.46	0.05
$\log_{10}\mathcal{B}$ (precession)	-	-0.68

Table 3.9: Bayes' Factors between: HOMs vs non-HOMs, non-preecssion vs zero spin and precessing spin vs non-precession

Impact of Higher Order Modes The dominant mode of gravitational radiation, $(2,2)$, radiates primarily in the direction of net angular momentum, while the HOMs carry radiation from off-axis asymmetry. Including the latter in the GW models breaks degeneracy on the extrinsic (particularly inclination angle, ι) as well as intrinsic binary parameters (particularly mass ratio, q). For the NRHybSur3dq8 model, we find that including HOMs narrows our estimate of the inclination angle of GW170502 by ∼25%, and more strongly excludes edge-on configurations (see Figure 3.26). While the $\ell = 2$ modes hinted at a low probability for an edge-on orientation ($\iota \sim 90^\circ$), including $\ell \leq 5$ completely rules it out within 90% confidence intervals; in fact, this analysis suggests that the binary is close to face-on. Since face-on configurations are more easily detected, the HOMs pushed the distance of GW170502 out by ∼10%. This increase in distance (redshift) directly translates into a lower mass for the BHs. For comparison, the median value of the primary BH mass using NRHybSur3dq8 is $m_1^{\rm src} = 94\ \mathrm{M}_\odot$ for $\ell \leq 5$, while $m_1^{\rm src} = 104\ \mathrm{M}_\odot$ for the $\ell = 2$ case.

While the inclusion of HOMs have a significant impact on the posteriors of GW170502, we do not find compelling evidence for their presence in the data. To quantify this, we computed the Bayes' Factor ($\mathcal{B}$) between the $\ell = 2$ and $\ell \leq 5$ ($\ell \leq 4$) case for both the waveform models. As stated in Table 3.9, we find $\log_{10}\mathcal{B}$ to be -0.03 and 0.00 for the NRHybSur3dq8 and NRSur7dq4 signal models respectively.

Discussion Our study demonstrates the necessary combination of parameter inference and waveform modeling techniques to constrain IMBH binary mergers. We apply this machinery to GW170502, the heaviest and loudest BBH trigger found in Advanced LIGO between 2015-2017. Using the most sophisticated GW models, we find that the primary and secondary BH masses of this trigger would correspond to ∼90 $\mathrm{M}_\odot$ and ∼ 60 $\mathrm{M}_\odot$. While not reflected in the

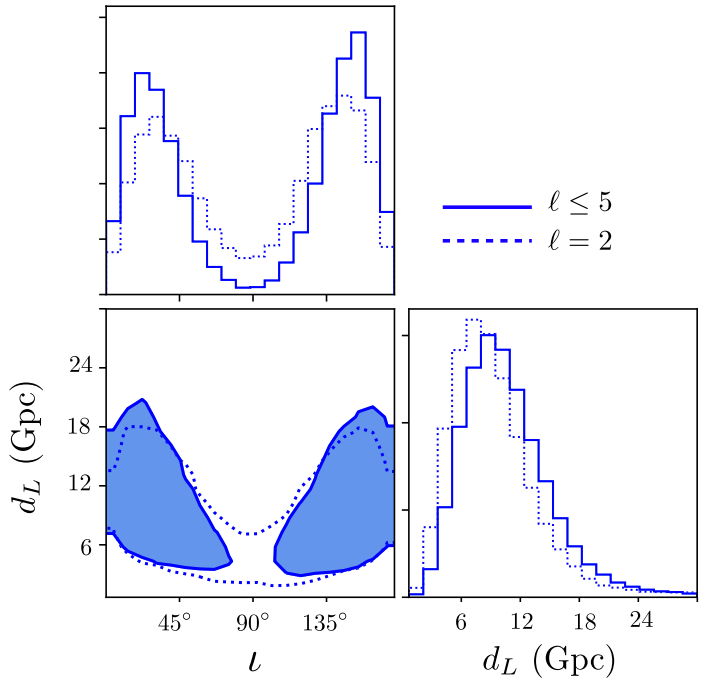

Figure 3.26: Marginalized posteriors of inclination (ι) and luminosity distance (d_L) for GW170502 using `RIFT`. Two-dimensional contours show 90% intervals for `NRHybSur3dq8` model with (solid line) and without (dotted line) including higher order modes of gravitational radiation.

Bayes' Factors, there is noticeable shift in the posteriors using the HOMs and $\ell = 2$ modes. It narrows the constraints on inclination and distance, thus reducing the uncertainty in BH masses.

In the next era of GW astronomy, GW170502-like events would be detected in mutliband network of earth-based detectors and space missions such as LISA and the deci-Hz Observatory [180, 181, 182]. This will increase the detection confidence of such sub-threshold triggers, and substantially improve the constraints on the masses and spins of the two BHs [183, 184]. For a future study, we will extend our machinery to explore the impact on parameter estimation of using fully general relativistic simulations of BBH coalescence.

Chapter 4

Analysis of O3a Events

4.1 Preamble

In the previous chapter, we demonstrated that the impact of waveform and HOM systematics can have a major impact on the parameters recovered. This can in turn affect other analysis dependent on accurate PE such as tests of general relativity, constraints on nuclear matter, and population inference. In this chapter, we now demonstrate how these systematics can impact analysis of real GW data. All the work presented here will eventually be published in exceptional event papers [185, 6] or in the O3a Catalog paper in the near future [7]. Below is a list summarizing all the results contained in this chapter:

- I analyzed a subset of the events detected in the first half of O3 with 15 binary black hole candidates and one binary neutron star candidate: GW190425, GW190513, GW190517, GW190519, GW190521g, GW190521r, GW190602, GW190620, GW190630, GW190701, GW190706, GW190719, GW190828j, GW190828ℓ, GW190909, GW190929 (see Figures 4.2-4.35)

- These results were run multiple times with different waveform models; the models that were used in this chapter include: SEOBNRv4P, SEOBNRv4PHM, TEOBResumS, SEOBNRv4T_surrogate, NR and NRSur7dq4; see Section 1.2

- These models were individually ran multiple times with varying degrees of HOMs.

SEOBNRv4P and SEOBNRv4PHM were used on all the events, TEOBResumS and SEOBNRv4T_surrogate were used for GW190425 with a low spin and high spin prior. NRSur7dq4 was used on all the high mass events with one $\ell_{\rm max} = 2$ analysis and one HOM analysis. NR was used only on GW190521g.

- The GW190521g analyses, which appeared in part in [6], includes SEOBNRv4PHM, NRSur7dq4, and NR compared directly to the data (see Section 1.3.2 for a brief summary and see [18] for a more detail description and validation)

- To quantify the information gained when including precession or HOMs, I calculated the Bayes Factors for all the binary BHs. GW190519 had the highest Bayes Factor when including HOMs with ~ 15, and GW190521g had the highest Bayes Factor when including precession with ~ 4.4. See Table 4.2.

- As shown in Chapter 3, the impact of waveform and HOM systematics is significant in many of the results presented here. Waveform systematics had the biggest impact on GW190521r with a JSD between IMRPhenomPv2 and SEOBNRv4P $\chi_{\rm eff}$ distributions of ~ 0.106 (see Figure 4.37). HOM systematics had the biggest impact on GW190519 with a Bayes Factor of ~ 15 and favored more asymmetric mass ratios, edge on orientations, lower spins, and smaller distances.

I emphasize that while the results presented are the analyses I did personally with RIFT, many more analyses were done with RIFT on other events that will appear in [7]. All the SEOBNRv4P, SEOBNRv4PHM, TEOBResumS, SEOBNRv4T_surrogate, and NRSur7dq4 results in this chapter are runs I did; however, all the IMRPhenomPv2 and IMRPhenomPv2_NRTidal_ROQ results presented in Figures 4.2-4.35 are done by the LVC for the O3a Catalog paper [7]. I do not claim these results as my own, but merely use them as a comparison for my own results.

4.2 Sources Properties of the LVC's O3a Events

In this section, I present 16 PE analyses done using RIFT on a subset of the first half of events in the LVC's 3rd observing Run. In general, we assume a low-frequency cutoff of $f_{min} = 20$Hz; however, we use a $f_{min} = 11$Hz for GW190521g and $f_{min} = 19.4$Hz for GW190425. The f_{max} was chosen depending on on the size of the PSDs generated for each event, either 255.75, 511.75, or 1023.75 Hz while the sample rate was defaulted to $s_{rate} = 4096$Hz. These PSDs characterizing the noise at the times of each of the events was generated by BayesWave using the segment length of data used in the inference [186, 187]. For most of the events, we analyzed 4 seconds of data. For GW190630, GW190828ℓ and the SEOBNRv4PHM run for GW190521g, we analyzed 8 seconds of data. As with the priors in [185, 6, 7], we assuming uniform priors in spin magnitudes and in the cube of the luminosity distance. For the mass parameters, the prior is uniform in the detector-frame. For a more detailed description on of data handling, searches, PE settings, etc., please see GWTC-2 [7].

4.2.1 Masses

The list of candidates presented in Table 4.1 includes a wide range of total masses. The most massive binaries with their median total mass $M^{source} > 100M_{\odot}$ in decreasing order are: GW190521g given in the *detector* frame total mass M^{det} (though it is still highest in source frame as well, see [185, 6] for more details) (SEOBNRv4PHM: $276.3^{+41.2}_{-32.6}$; NRSur7dq4: $271.4^{+22.5}_{-24.7}$; NR: $270.5^{+23.1}_{-24.5}$; NR+NRSur7dq4: $270.9^{+23.5}_{-25.9}$), GW190602 (SEOBNRv4PHM: $115.1^{+17.9}_{-15.0}$; NRSur7dq4 all modes: $111.2^{+20.3}_{-15.4}$), GW190519 (SEOBNRv4PHM: $105.8^{+13.7}_{-12.7}$; NRSur7dq4 all modes: $101.1^{+16.5}_{-13.9}$), and GW190706 (SEOBNRv4PHM: $103.7^{+16.6}_{-12.8}$) putting them higher than any detected in O1/O2 Observing Runs as well as above several proposed limits introduced by pair-instability SN (PISN) [188, 15]. Other notable high mass systems with $M \sim 100M_{\odot}$ are: GW190701 (SEOBNRv4PHM: $95.1^{+11.9}_{-9.6}$; NRSur7dq4 all modes: $93.1^{+10.8}_{-8.8}$ and GW190929 (SEOBNRv4PHM: $90.4^{+20.7}_{-14.3}$). While many of the mass ratios presented here peak towards more unequal mass systems, they are still consistent with equal mass mergers. If higher order modes are important for a given candidate, the mass ratio is

one of the parameters most effected. For a more detail discussion on the impact of higher order modes, see Section 4.2.4. The odd-numbered figures between Figures 4.8-4.34 show the distributions for the different source frame mass parameters.

The least massive binary is GW190425 (SEOBNRv4T_surrogate low spin: $3.3^{+0.06}_{-0.05}$, high spin: $3.3^{+0.2}_{-0.07}$; TEOBResumS low spin: $3.3^{+0.06}_{-0.05}$, high spin: $3.3^{+0.2}_{-0.08}$). Given the inferred individual masses of the system, it is most likely a binary neutron star although the tidal constraints do not rule out a binary black hole or neutron star-black hole binary. Similar to GW170817[103], the source masses presented here are consistent with masses of known neutron stars. However, the total mass is greater than that of other binary neutron star systems detected. Figures 4.2 and 4.5 shows the mass parameters of the low and high spin priors. In both scenarios, the orange and blue represent the results done with RIFT using the non-precessing models SEOBNRv4T_surrogate and TEOBResumS and the green represents the results done with LALInference using IMRPhenomPv2_NRTidal_ROQ. For the low spin prior limit, the EOB results produce similar results presented in the GW190425 discovery paper [189] including both precessing IMRPhenomPv2_NRTidal_ROQ and non-precessing IMRPhenomD_NRTidal_ROQ (Not shown in Figures 4.2 and 4.5). When allowing for larger spins, the IMRPhenomD_NRTidal_ROQ and EOB have modest differences with the individual masses. Since there is more flexibility and configurations with the higher spin limitation with the precessing model IMRPhenomPv2_NRTidal_ROQ, there is more support for larger/smaller m_1/m_2 mass values. Similarly we see largely consistent results between the different models for the low spin limit; however, we see the precessing model produce a more asymmetric distribution compare to the nonprecessing models when allowing for higher spin configurations.

4.2.2 Spins

As in previous chapters, we focus on constraining the mass-weighted aligned spin (the effective spin) $\chi_{\rm eff}$. The even-numbered figures between Figures 4.9-4.35 includes the effective spin distributions for the different candidates. Most of the binaries have spin magnitudes consistent

with zero within uncertainties; however, some candidates can be constrained away from zero. At a 90% credibility, we find in decreasing (based off the SEOBNRv4PHM result) median effective spin: GW190517 (SEOBNRv4PHM: $0.56^{+0.19}_{-0.18}$; NRSur7dq4 (all modes): $0.50^{+0.18}_{-0.19}$), GW190719 (SEOBNRv4PHM: $0.38^{+0.27}_{-0.33}$), GW190620 (SEOBNRv4PHM: $0.37^{+0.21}_{-0.24}$; NRSur7dq4 (all modes): $0.31^{+0.21}_{-0.24}$), GW190519 (SEOBNRv4PHM: $0.35^{+0.19}_{-0.24}$; NRSur7dq4 (all modes): $0.32^{+0.18}_{-0.21}$), GW190706 (SEOBNRv4PHM: $0.32^{+0.25}_{-0.30}$; NRSur7dq4 (all modes): $0.33^{+0.24}_{-0.29}$), and GW190828j (SEOBNRv4PHM: $0.22^{+0.14}_{-0.16}$; NRSur7dq4 (all modes): $0.18^{+0.13}_{-0.16}$). While we find no system among the events I analyzed is confidently $\chi_{\rm eff} < 0$, the system with the most negative median effective spin was GW190701 (SEOBNRv4PHM: $-0.04^{+0.23}_{-0.30}$; NRSur7dq4 (all modes): $-0.08^{+0.22}_{-0.27}$).

To estimate the precession of a given system, the even-numbered between Figures 4.9-4.35 also include results showing the effective precession spin χ_p [46]. For most of these posterior distributions, the results are similar to the assumed prior. To quantify any evidence of precession, I calculate a Bayes Factors in favor of precession for all the candidates presented in Table 4.2. In the same vein as the χ_p distributions being similar to the prior, most of the Bayes Factors are consistent with 1 (neither favoring precessing nor non-precessing systems). GW190521g and GW190828ℓ gives some mild evidence of precession with corresponding Bayes Factors of 4.38 and 2.28 respectively.

4.2.3 Distance and θ_{jn}

The most distant candidates are most likely GW190909 (SEOBNRv4PHM: $d_L = 4.87^{+3.71}_{-2.55}$) and GW190706 (SEOBNRv4PHM: $d_L = 4.78^{+2.26}_{-1.82}$; NRSur7dq4: $d_L = 5.55^{+2.55}_{-2.47}$). These along with other events are larger than anything detected from O2 [15]. The nearest event was GW190425 where using SEOBNRv4T_surrogate yields a distance of $0.16^{+0.069}_{-0.073}$ (low spin) and $0.17^{+0.076}_{-0.075}$ (high spin) and TEOBResumS yields a distance of $0.16^{+0.068}_{-0.073}$ (low spin) and $0.17^{+0.074}_{-0.075}$ (high spin). This is 4 times farther away than the binary neutron star system GW170817 found in O2 [15]. While the individual distance values may not be as interesting, the effects of HOMs on the distance (and therefore the source frame masses) distribution

can be significant. In my analyses, HOMs have affected the distance distributions to varying degrees, see Figures 4.9-4.35

Most of the events presented here have a bimodal θ_{jn} distribution that are likely face on orientations but include a non-zero probability of being edge on as well. While GW190701 is also likely face on, Figure 4.23 shows that it favors a particular orientation (i.e. non-bimodal). When including HOMs in the analysis, Figure 4.13 shows GW190519 favors more edge on orientations. To a lesser degree, Figure 4.21 shows GW190630 more confidently allows for edge on cases. See Section 4.2.4 for more details about the impact of HOMs.

4.2.4 Impact of higher order modes

Figure 4.36 shows the results of four events (GW190519, GW190602, GW190706, and GW190929) where the impact of HOMs was most dramatic. In the figure, I use two models which include some number of HOMs (SEOBNRv4PHM and NRSur7dq4; see Section 1.2 for details on which modes are included) and one model which only include the dominant (2,2) mode (SEOBNRv4P). The most significant impact was on GW190519. When including HOMs, the parameters inferred favored more asymmetric mass ratios and thus a higher source-frame primary mass, edge on orientations, lower spins, and smaller distances. To quantify this the information gained when including HOMs, I calculated a Bayes Factors~ 15. As seen in Figures 4.20;4.21 and Figures 4.30;4.31, HOMs also have a noticeable impact on the analyses for GW190630 and GW190828ℓ.

4.2.5 Impact of waveform systematics

Similar to calculations in Section 3.4.3.4 with Eq. 3.4.7, we can calculated the JSD to quantify waveform systematics between different results (see Section 3.4.3.4 for more details). The largest JSD for the events presented here between the two (2,2) only models: IMRPhenomPv2 and SEOBNRv4P is 0.103 for the $\chi_{\rm eff}$ for GW190521r. This is illustrated in Figure 4.37 where the green and orange distributions are representing the SEOBNRv4P and IMRPhenomPv2 results respectively. Here SEOBNRv4P prefers more positive effective spin;

however, both SEOBNRv4P and IMRPhenomPv2 are still consistent with $\chi_{\rm eff} = 0$ configurations. There are also more minor differences in mass ratio and luminosity distance. While the other events include various degrees of differences between the models, there is always substantial overlap between the different models posteriors (that include a similar amount of modes). While there is visible systematic waveform differences present in the posteriors, they are still secondary to statistical effects.

4.2.6 Analysis of GW190521g

In this section, we emphasize the analysis done for GW190521g. Figure 4.1 shows the detector frame masses for not only SEOBNRv4PHM and NRSur7dq4 but also analyses that directly compares NR waveforms to the data and a combined-grind NR+NRSur7dq4 result (see Section 2.4.1 for more details). This event has the largest masses ever detected with a detector-frame total mass of $270.5^{+23.1}_{-24.5}$ (NR), $270.9^{+23.5}_{-25.9}$ (NR+NRSur7dq4, $271.4^{+22.5}_{-24.7}$ (NRSur7dq4, and $276.3^{+41.2}_{-32.6}$ (SEOBNRv4PHM) as well as mass ratios with slightly unequal masses $0.79^{+0.19}_{-0.26}$ (NR), $0.79^{+0.18}_{-0.28}$ (NR+NRSur7dq4), $0.82^{+0.16}_{-0.26}$ (NRSur7dq4), and $0.75^{+0.22}_{-0.34}$ (SEOBNRv4PHM). While the components of the spins aligned with the orbital angular momentum are centered around zero, there seems to be support for high transverse spins. Table 4.2 shows the Bayes Factor calculated when including precession to be ~ 4.38. As seen in Figure 4.1, NR has support for noticeable lower transverse spin configurations compared to the other parameters. This is most likely due to the sparse number of NR simulations with high transverse spins. For a more detailed description of parameters recovered, see Table 4.4 as well as [185, 6]

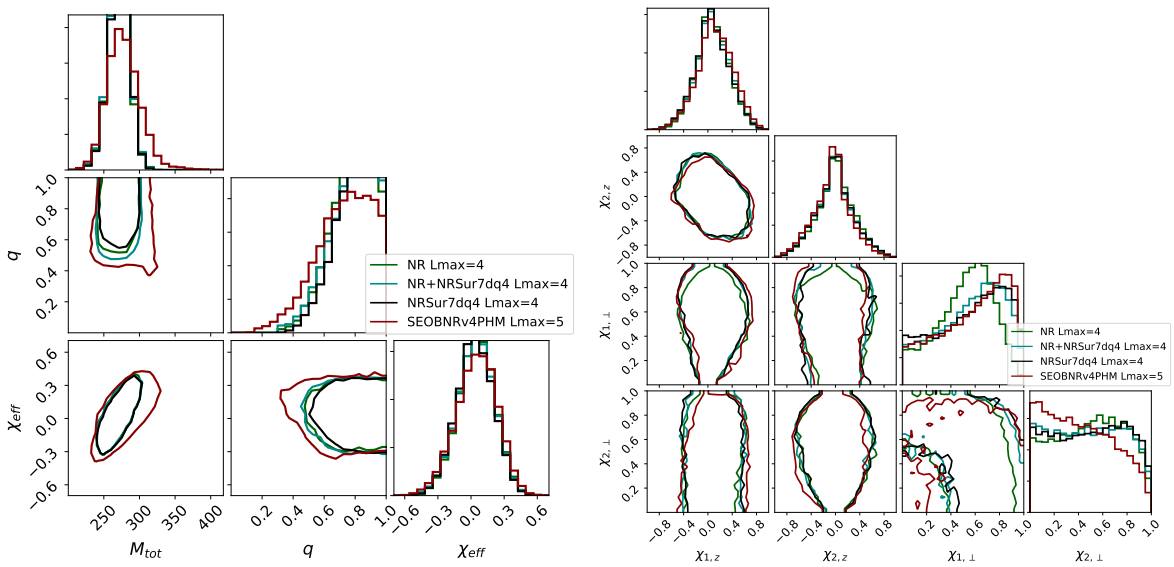

Figure 4.1: **GW190521g Binary Black Hole System:** This figure shows corner plots of both mass and spin parameters from different analyses of GW190521g. The results include models SEOBNRv4PHM (red), NRSur7dq4 (black) with all available modes, NR (green) including all $\ell_{\max} \leq 4$, and a combined grid of of the latter two results NR+NRSur7dq4 all done with RIFT. There are noticeable differences between different waveforms. NR has support for lower transverse spin magnitude configurations. This is most likely due to the fact that there are not as many NR simulations with high transverse spins. See Section 4.2.6 for a more detailed description.

Event	Model	$m_1^{source}(M_\odot)$	$m_2^{source}(M_\odot)$	$M^{source}(M_\odot)$	$\mathcal{M}^{source}(M_\odot)$	q	$\chi_{\rm eff}$	D_L(Gpc)
GW190513	SEOBNRv4P	$35.5^{+10.5}_{-8.8}$	$18.3^{+7.2}_{-4.4}$	$54.2^{+8.7}_{-6.3}$	$21.8^{+3.5}_{-2.1}$	$0.52^{+0.39}_{-0.20}$	$0.13^{+0.27}_{-0.18}$	$2.05^{+0.98}_{-0.89}$
	SEOBNRv4PHM	$36.0^{+9.6}_{-9.2}$	$17.9^{+7.4}_{-4.3}$	$54.0^{+8.5}_{-6.0}$	$21.5^{+3.7}_{-1.9}$	$0.50^{+0.41}_{-0.18}$	$0.13^{+0.30}_{-0.18}$	$2.12^{+0.95}_{-0.73}$
GW190517	SEOBNRv4P	$38.0^{+14.1}_{-8.6}$	$24.5^{+7.2}_{-6.9}$	$63.5^{+10.9}_{-10.1}$	$26.5^{+4.1}_{-4.0}$	$0.65^{+0.30}_{-0.29}$	$0.56^{+0.19}_{-0.19}$	$2.14^{+1.77}_{-1.01}$
	SEOBNRv4PHM	$36.7^{+11.5}_{-7.6}$	$25.5^{+6.7}_{-7.1}$	$63.0^{+9.7}_{-9.7}$	$26.5^{+4.0}_{-4.0}$	$0.70^{+0.26}_{-0.29}$	$0.56^{+0.19}_{-0.18}$	$2.13^{+1.77}_{-0.99}$
	NRSur7dq4 ($\ell = 2$)	$36.3^{+11.8}_{-7.6}$	$24.0^{+6.7}_{-6.5}$	$61.1^{+9.8}_{-8.9}$	$26.6^{+3.9}_{-3.6}$	$0.67^{+0.29}_{-0.27}$	$0.51^{+0.18}_{-0.18}$	$2.12^{+1.63}_{-0.98}$
	NRSur7dq4 (all modes)	$35.7^{+12.0}_{-7.6}$	$23.8^{+6.6}_{-6.7}$	$60.1^{+9.8}_{-8.9}$	$25.2^{+3.9}_{-3.7}$	$0.67^{+0.28}_{-0.29}$	$0.50^{+0.18}_{-0.19}$	$2.24^{+1.78}_{-1.07}$
GW190519	SEOBNRv4P	$60.4^{+15.7}_{-12.2}$	$40.8^{+11.3}_{-12.3}$	$101.0^{+16.2}_{-12.0}$	$42.4^{+7.3}_{-5.3}$	$0.69^{+0.27}_{-0.29}$	$0.39^{+0.18}_{-0.21}$	$3.67^{+1.67}_{-1.53}$
	SEOBNRv4PHM	$65.8^{+10.8}_{-11.3}$	$40.3^{+10.4}_{-10.2}$	$105.8^{+13.7}_{-12.7}$	$44.2^{+6.3}_{-6.2}$	$0.61^{+0.26}_{-0.18}$	$0.35^{+0.19}_{-0.24}$	$2.84^{+1.60}_{-1.01}$
	NRSur7dq4 ($\ell = 2$)	$57.4^{+17.1}_{-11.1}$	$39.7^{+10.9}_{-11.2}$	$96.8^{+17.2}_{-11.2}$	$40.8^{+7.5}_{-4.9}$	$0.70^{+0.26}_{-0.29}$	$0.35^{+0.18}_{-0.20}$	$3.89^{+1.64}_{-1.72}$
	NRSur7dq4 (all modes)	$62.1^{+12.0}_{-13.4}$	$39.5^{+11.7}_{-10.5}$	$101.1^{+16.5}_{-13.9}$	$42.3^{+7.6}_{-6.2}$	$0.63^{+0.30}_{-0.20}$	$0.32^{+0.18}_{-0.21}$	$3.12^{+2.18}_{-1.43}$
GW190521r	SEOBNRv4P	$43.1^{+7.2}_{-5.5}$	$31.9^{+5.7}_{-5.9}$	$74.8^{+7.1}_{-4.7}$	$32.0^{+3.2}_{-2.2}$	$0.74^{+0.22}_{-0.20}$	$0.11^{+0.09}_{-0.14}$	$1.31^{+0.36}_{-0.58}$
	SEOBNRv4PHM	$42.3^{+5.5}_{-5.1}$	$32.3^{+5.3}_{-5.2}$	$74.4^{+6.6}_{-4.2}$	$31.9^{+3.0}_{-2.0}$	$0.76^{+0.20}_{-0.18}$	$0.12^{+0.09}_{-0.12}$	$1.34^{+0.33}_{-0.55}$
	NRSur7dq4 ($\ell = 2$)	$42.9^{+8.6}_{-5.7}$	$31.6^{+5.8}_{-7.4}$	$74.3^{+6.9}_{-4.8}$	$31.7^{+3.1}_{-2.4}$	$0.74^{+0.23}_{-0.26}$	$0.05^{+0.11}_{-0.13}$	$1.20^{+0.40}_{-0.53}$
	NRSur7dq4 (all modes)	$41.9^{+5.8}_{-4.8}$	$32.6^{+5.2}_{-6.4}$	$74.1^{+6.9}_{-4.6}$	$31.9^{+3.1}_{-2.5}$	$0.78^{+0.19}_{-0.20}$	$0.08^{+0.09}_{-0.13}$	$1.25^{+0.41}_{-0.55}$
GW190602	SEOBNRv4P	$72.7^{+20.9}_{-16.0}$	$43.5^{+16.1}_{-18.2}$	$116.1^{+19.0}_{-15.5}$	$47.7^{+9.2}_{-9.1}$	$0.60^{+0.35}_{-0.31}$	$0.13^{+0.24}_{-0.24}$	$2.68^{+1.83}_{-1.22}$
	SEOBNRv4PHM	$67.9^{+16.2}_{-12.5}$	$47.8^{+13.2}_{-16.9}$	$115.1^{+17.9}_{-15.0}$	$48.7^{+8.3}_{-8.1}$	$0.71^{+0.25}_{-0.32}$	$0.11^{+0.26}_{-0.26}$	$2.92^{+1.84}_{-1.16}$
	NRSur7dq4 ($\ell = 2$)	$71.8^{+19.7}_{-16.3}$	$41.0^{+18.9}_{-15.3}$	$113.4^{+20.8}_{-15.6}$	$46.2^{+10.7}_{-8.0}$	$0.58^{+0.37}_{-0.28}$	$0.10^{+0.24}_{-0.24}$	$2.77^{+2.05}_{-1.32}$
	NRSur7dq4 (all modes)	$65.9^{+16.5}_{-12.4}$	$46.3^{+14.4}_{-17.5}$	$111.2^{+20.3}_{-15.4}$	$46.9^{+9.7}_{-8.1}$	$0.72^{+0.25}_{-0.35}$	$0.09^{+0.24}_{-0.24}$	$3.26^{+2.25}_{-1.44}$
GW190620	SEOBNRv4P	$57.1^{+17.9}_{-12.6}$	$35.9^{+11.5}_{-11.8}$	$93.2^{+16.4}_{-12.6}$	$38.7^{+7.1}_{-5.8}$	$0.63^{+0.32}_{-0.28}$	$0.35^{+0.20}_{-0.25}$	$2.97^{+1.66}_{-1.34}$
	SEOBNRv4PHM	$56.0^{+16.1}_{-11.7}$	$36.8^{+11.0}_{-12.2}$	$92.5^{+16.3}_{-12.1}$	$38.6^{+7.4}_{-5.8}$	$0.66^{+0.30}_{-0.29}$	$0.37^{+0.21}_{-0.24}$	$3.12^{+1.56}_{-1.32}$
	NRSur7dq4 ($\ell = 2$)	$55.6^{+16.3}_{-12.6}$	$33.4^{+11.5}_{-10.8}$	$89.0^{+15.5}_{-11.3}$	$36.7^{+6.9}_{-5.2}$	$0.60^{+0.34}_{-0.27}$	$0.28^{+0.22}_{-0.24}$	$3.02^{+1.64}_{-1.44}$
	NRSur7dq4 (all modes)	$54.3^{+15.1}_{-11.9}$	$33.3^{+11.1}_{-10.1}$	$87.3^{+16.9}_{-10.7}$	$36.2^{+7.5}_{-4.9}$	$0.61^{+0.33}_{-0.25}$	$0.31^{+0.21}_{-0.24}$	$3.33^{+1.69}_{-1.57}$
GW190630	SEOBNRv4P	$35.9^{+8.1}_{-6.7}$	$22.3^{+5.5}_{-4.3}$	$58.4^{+5.5}_{-4.6}$	$24.4^{+2.1}_{-1.7}$	$0.62^{+0.31}_{-0.20}$	$0.10^{+0.13}_{-0.13}$	$1.02^{+0.49}_{-0.44}$
	SEOBNRv4PHM	$35.1^{+6.9}_{-5.9}$	$23.3^{+5.2}_{-4.8}$	$58.7^{+4.9}_{-4.6}$	$24.7^{+2.2}_{-1.9}$	$0.66^{+0.28}_{-0.21}$	$0.11^{+0.13}_{-0.13}$	$0.97^{+0.52}_{-0.40}$
	NRSur7dq4 ($\ell = 2$)	$35.4^{+7.7}_{-6.4}$	$22.6^{+5.3}_{-4.4}$	$58.1^{+5.5}_{-4.5}$	$24.4^{+2.1}_{-1.8}$	$0.64^{+0.30}_{-0.20}$	$0.08^{+0.12}_{-0.13}$	$1.01^{+0.49}_{-0.44}$
	NRSur7dq4 (all modes)	$34.7^{+6.8}_{-5.5}$	$23.6^{+5.2}_{-5.1}$	$58.6^{+4.6}_{-4.8}$	$24.7^{+2.1}_{-2.0}$	$0.68^{+0.27}_{-0.22}$	$0.09^{+0.11}_{-0.13}$	$0.95^{+0.56}_{-0.42}$
GW190701	SEOBNRv4P	$55.7^{+14.1}_{-9.1}$	$40.5^{+9.1}_{-12.3}$	$96.0^{+12.8}_{-10.0}$	$40.8^{+5.6}_{-5.1}$	$0.74^{+0.23}_{-0.31}$	$-0.04^{+0.33}_{-0.29}$	$2.05^{+0.76}_{-0.76}$
	SEOBNRv4PHM	$54.1^{+12.1}_{-8.0}$	$41.3^{+8.3}_{-11.7}$	$95.1^{+11.9}_{-9.6}$	$40.6^{+5.3}_{-4.8}$	$0.77^{+0.20}_{-0.31}$	$-0.04^{+0.23}_{-0.30}$	$2.14^{+0.74}_{-0.69}$
	NRSur7dq4 ($\ell = 2$)	$55.1^{+13.7}_{-9.0}$	$39.5^{+9.0}_{-11.7}$	$94.4^{+11.9}_{-9.4}$	$40.0^{+5.4}_{-4.7}$	$0.73^{+0.24}_{-0.30}$	$-0.08^{+0.22}_{-0.27}$	$1.98^{+0.80}_{-0.75}$
	NRSur7dq4 (all modes)	$53.1^{+11.2}_{-7.7}$	$40.3^{+8.1}_{-11.1}$	$93.1^{+10.8}_{-8.8}$	$39.8^{+4.9}_{-4.4}$	$0.76^{+0.21}_{-0.29}$	$-0.08^{+0.22}_{-0.27}$	$2.17^{+0.83}_{-0.78}$
GW190706	SEOBNRv4P	$63.7^{+22.0}_{-14.4}$	$40.3^{+13.2}_{-14.2}$	$103.7^{+22.0}_{-13.7}$	$43.0^{+9.5}_{-6.3}$	$0.64^{+0.32}_{-0.30}$	$0.41^{+0.22}_{-0.28}$	$5.23^{+2.24}_{-2.32}$
	SEOBNRv4PHM	$66.3^{+14.0}_{-15.1}$	$38.0^{+12.9}_{-11.7}$	$103.7^{+16.6}_{-12.8}$	$42.7^{+8.0}_{-6.3}$	$0.58^{+0.33}_{-0.22}$	$0.32^{+0.25}_{-0.30}$	$4.78^{+2.26}_{-1.82}$
	NRSur7dq4 ($\ell = 2$)	$61.3^{+22.2}_{-13.2}$	$41.2^{+13.0}_{-14.3}$	$102.0^{+23.6}_{-13.4}$	$42.7^{+10.0}_{-6.0}$	$0.68^{+0.28}_{-0.34}$	$0.39^{+0.22}_{-0.25}$	$5.34^{+2.29}_{-2.47}$
	NRSur7dq4 (all modes)	$60.7^{+16.0}_{-13.7}$	$38.9^{+11.9}_{-12.9}$	$98.9^{+18.4}_{-12.5}$	$41.2^{+8.3}_{-5.8}$	$0.65^{+0.30}_{-0.28}$	$0.33^{+0.24}_{-0.29}$	$5.55^{+2.55}_{-2.47}$
GW190719	SEOBNRv4P	$39.0^{+21.9}_{-12.6}$	$20.0^{+8.5}_{-6.4}$	$59.4^{+20.9}_{-11.8}$	$23.7^{+6.7}_{-4.0}$	$0.51^{+0.42}_{-0.25}$	$0.39^{+0.25}_{-0.33}$	$4.30^{+2.82}_{-2.21}$
	SEOBNRv4PHM	$36.1^{+18.7}_{-10.3}$	$20.5^{+8.4}_{-6.9}$	$57.1^{+17.8}_{-10.3}$	$23.2^{+6.1}_{-3.7}$	$0.58^{+0.37}_{-0.30}$	$0.38^{+0.27}_{-0.33}$	$4.56^{+2.79}_{-2.19}$
GW190828j	SEOBNRv4P	$32.5^{+6.9}_{-4.2}$	$25.6^{+4.3}_{-4.9}$	$58.0^{+7.1}_{-4.3}$	$24.9^{+3.1}_{-1.9}$	$0.80^{+0.18}_{-0.24}$	$0.21^{+0.13}_{-0.16}$	$2.24^{+0.59}_{-0.92}$
	SEOBNRv4PHM	$31.9^{+5.9}_{-3.8}$	$26.1^{+4.3}_{-4.6}$	$57.8^{+6.9}_{-4.4}$	$24.9^{+3.1}_{-1.9}$	$0.83^{+0.15}_{-0.23}$	$0.22^{+0.14}_{-0.16}$	$2.28^{+0.59}_{-0.90}$
	NRSur7dq4 ($\ell = 2$)	$32.1^{+6.5}_{-4.0}$	$25.6^{+4.2}_{-4.8}$	$57.6^{+7.1}_{-4.4}$	$24.7^{+3.0}_{-2.0}$	$0.81^{+0.17}_{-0.24}$	$0.17^{+0.13}_{-0.16}$	$2.08^{+0.63}_{-0.85}$
	NRSur7dq4 (all modes)	$31.6^{+5.6}_{-3.9}$	$25.7^{+4.4}_{-4.5}$	$57.0^{+7.6}_{-4.3}$	$24.5^{+3.4}_{-1.9}$	$0.82^{+0.16}_{-0.21}$	$0.18^{+0.13}_{-0.16}$	$2.20^{+0.64}_{-0.95}$
GW190828ℓ	SEOBNRv4P	$22.2^{+9.0}_{-6.4}$	$10.5^{+3.7}_{-2.5}$	$32.9^{+6.8}_{-3.9}$	$13.1^{+1.3}_{-0.9}$	$0.47^{+0.42}_{-0.21}$	$0.05^{+0.17}_{-0.16}$	$1.76^{+0.64}_{-0.72}$
	SEOBNRv4PHM	$23.7^{+7.2}_{-7.1}$	$10.2^{+3.6}_{-2.1}$	$34.0^{+5.5}_{-4.4}$	$13.2^{+1.2}_{-0.9}$	$0.43^{+0.39}_{-0.16}$	$0.08^{+0.16}_{-0.16}$	$1.68^{+0.62}_{-0.61}$
GW190909	SEOBNRv4P	$41.2^{+22.3}_{-10.5}$	$27.1^{+10.9}_{-10.4}$	$68.7^{+23.6}_{-13.1}$	$28.5^{+9.3}_{-5.5}$	$0.68^{+0.28}_{-0.38}$	$-0.06^{+0.32}_{-0.36}$	$4.64^{+3.38}_{-2.57}$
	SEOBNRv4PHM	$42.9^{+50.3}_{-11.9}$	$27.9^{+13.2}_{-11.0}$	$70.8^{+54.9}_{-14.8}$	$29.4^{+18.2}_{-6.3}$	$0.64^{+0.31}_{-0.38}$	$-0.03^{+0.45}_{-0.36}$	$4.87^{+3.71}_{-2.55}$
GW190929	SEOBNRv4P	$60.8^{+44.1}_{-19.4}$	$29.6^{+15.3}_{-13.6}$	$90.3^{+44.4}_{-16.4}$	$35.2^{+14.7}_{-6.8}$	$0.48^{+0.45}_{-0.30}$	$0.08^{+0.29}_{-0.30}$	$4.23^{+3.42}_{-2.61}$
	SEOBNRv4PHM	$64.4^{+21.8}_{-19.1}$	$25.8^{+14.3}_{-9.6}$	$90.4^{+20.7}_{-14.3}$	$34.3^{+8.6}_{-6.3}$	$0.40^{+0.41}_{-0.19}$	$0.04^{+0.27}_{-0.27}$	$3.75^{+3.01}_{-1.68}$

Table 4.1: Median and 90% confidence intervals on source parameters of all binary black hole O3a events discussed in Ch. 4.

Event	Bayes Factor (HOM vs non-HOM)	Bayes Factor (precession vs non-precession)
GW190513	0.99	0.57
GW190517	1.03	0.95
GW190519	15.02	1.13
GW190521g	0.45*	4.38
GW190521r	1.17	0.88
GW190602	0.76	0.95
GW190620	0.99	0.75
GW190630	0.93	1.11
GW190701	0.72	0.94
GW190706	2.35	0.62
GW190719	0.60	0.84
GW190828j	0.87	0.82
GW190828ℓ	2.21	2.28
GW190909	0.74	1.39
GW190929	2.99	1.34

Table 4.2: Bayes Factors between higher order modes vs non-higher mode as well as between precession vs non-precession results. The (*) denotes the Bayes Factor calculated using NRSur7dq4 analyses instead of between SEOBNRv4P and SEOBNRv4PHM.

Model	Spin Limit	$m_1^{source}(M_\odot)$	$m_2^{source}(M_\odot)$	$M^{source}(M_\odot)$	$\mathcal{M}^{source}(M_\odot)$	q	$\chi_{\rm eff}$	D_L(Gpc)	λ_1	λ_2	$\tilde{\Lambda}$
SEOBNRv4T_surrogate	low spin	$1.7^{+0.2}_{-0.09}$	$1.6^{+0.08}_{-0.1}$	$3.3^{+0.06}_{-0.05}$	$1.4^{+0.02}_{-0.02}$	$0.89^{+0.09}_{-0.15}$	$0.01^{+0.02}_{-0.01}$	$0.16^{+0.069}_{-0.073}$	303^{+1143}_{-280}	457^{+1652}_{-422}	432^{+999}_{-329}
	high spin	$1.9^{+0.5}_{-0.20}$	$1.5^{+0.2}_{-0.3}$	$3.3^{+0.2}_{-0.07}$	$1.4^{+0.02}_{-0.02}$	$0.78^{+0.19}_{-0.27}$	$0.03^{+0.07}_{-0.03}$	$0.17^{+0.076}_{-0.075}$	382^{+1803}_{-354}	891^{+2679}_{-825}	638^{+1504}_{-503}
TEOBResumS	low spin	$1.8^{+0.2}_{-0.09}$	$1.6^{+0.09}_{-0.1}$	$3.3^{+0.06}_{-0.05}$	$1.4^{+0.02}_{-0.02}$	$0.89^{+0.10}_{-0.15}$	$0.01^{+0.02}_{-0.01}$	$0.16^{+0.068}_{-0.073}$	285^{+1091}_{-263}	445^{+1637}_{-410}	413^{+991}_{-310}
	high spin	$1.9^{+0.5}_{-0.2}$	$1.4^{+0.2}_{-0.3}$	$3.3^{+0.2}_{-0.08}$	$1.4^{+0.02}_{-0.02}$	$0.75^{+0.22}_{-0.27}$	$0.04^{+0.06}_{-0.03}$	$0.17^{+0.074}_{-0.075}$	330^{+1549}_{-308}	904^{+3491}_{-833}	587^{+1546}_{-456}

Table 4.3: Median and 90% confidence intervals on source parameters of GW190425 discussed in Ch. 4.

Event	Model	$m_1^{det}(M_\odot)$	$m_2^{det}(M_\odot)$	$M^{det}(M_\odot)$	$\mathcal{M}^{det}(M_\odot)$	q	$\chi_{\rm eff}$	$\chi_{\rm p}$
GW190521g	NR	$151.1^{+28.5}_{-17.3}$	$119.1^{+18.8}_{-29.0}$	$270.5^{+23.1}_{-24.5}$	$116.2^{+10.1}_{-12.2}$	$0.79^{+0.19}_{-0.26}$	$0.05^{+0.26}_{-0.29}$	$0.62^{+0.25}_{-0.30}$
	NR+NRSur7dq4	$150.9^{+29.8}_{-18.1}$	$119.7^{+18.6}_{-31.5}$	$270.9^{+23.5}_{-25.9}$	$116.3^{+10.4}_{-13.4}$	$0.79^{+0.18}_{-0.28}$	$0.05^{+0.26}_{-0.29}$	$0.68^{+0.25}_{-0.36}$
	NRSur7dq4	$149.4^{+16.9}_{-17.0}$	$121.8^{+16.9}_{-27.6}$	$271.4^{+22.5}_{-24.7}$	$117.0^{+9.8}_{-12.1}$	$0.82^{+0.16}_{-0.26}$	$0.04^{+0.27}_{-0.28}$	$0.70^{+0.25}_{-0.38}$
	SEOBNRv4PHM	$157.1^{+58.4}_{-22.3}$	$117.4^{+24.5}_{-37.1}$	$276.3^{+41.2}_{-32.6}$	$117.8^{+14.5}_{-16.8}$	$0.75^{+0.22}_{-0.34}$	$0.06^{+0.28}_{-0.33}$	$0.68^{+0.26}_{-0.42}$

Table 4.4: Median and 90% confidence intervals on source parameters of GW190521g discussed in Section 4.2.6.

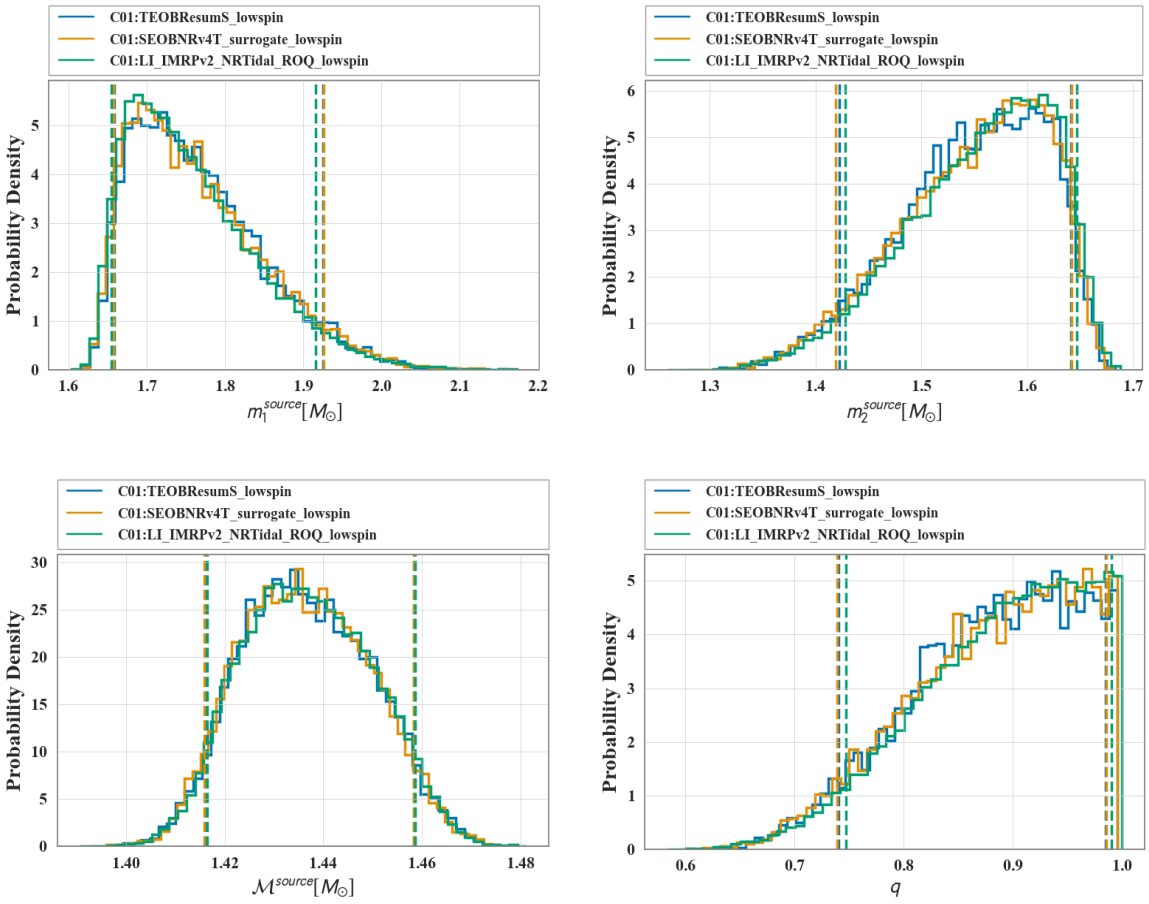

Figure 4.2: **GW190425 binary neutron star candidate with a low spin prior I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190425 using a low spin prior limit ($\chi_{i,z} \in [-0.05, 0.05]$). The results include models TEOBResumS (blue) and SEOBNRv4T_surrogate (orange) done with RIFT and IMRPhenomPv2_NRTidal_ROQ (green) done with LALInference. The results are largely consistent across all waveforms.

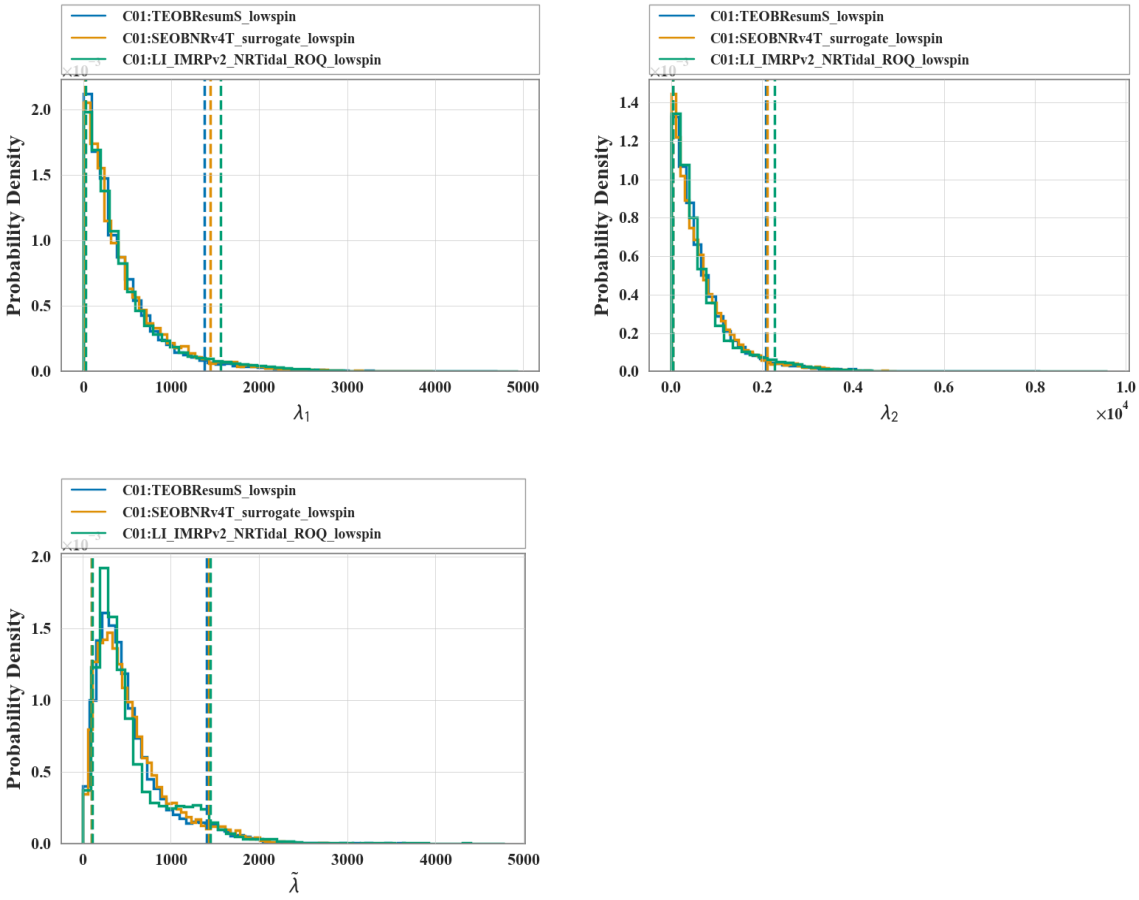

Figure 4.3: **GW190425 binary neutron star candidate with a low spin prior II:** This figure shows panels of the tidal parameters from the different analyses of GW190425 using a low spin prior limit ($\chi_{i,z} \in [-0.05, 0.05]$). The results include models TEOBResumS (blue) and SEOBNRv4T_surrogate (orange) done with RIFT and IMRPhenomPv2_NRTidal_ROQ (green) done with LALInference. The results are largely consistent across all waveforms.

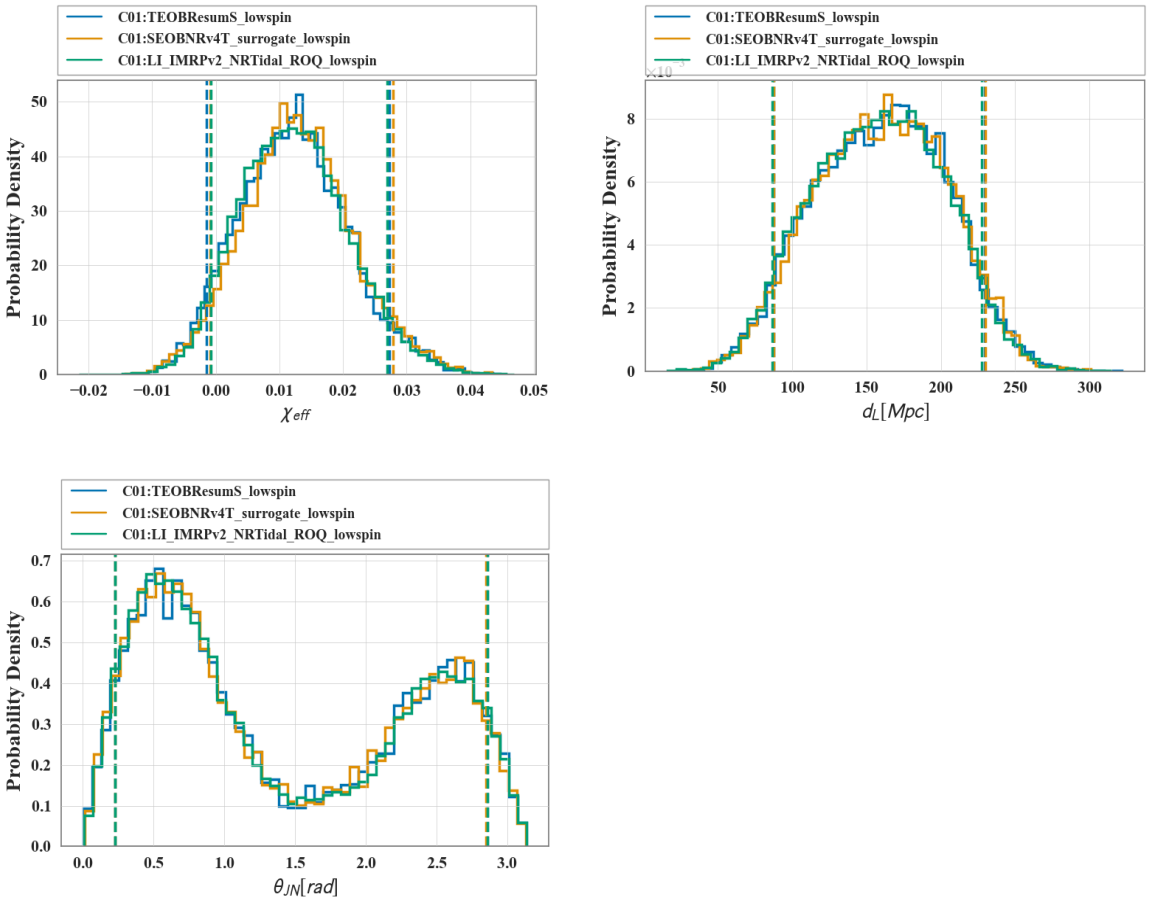

Figure 4.4: **GW190425 binary neutron star candidate with a low spin prior III:** This figure shows panels of the 1D marginal distributions for the $\chi_{\rm eff}$, luminosity distance (d_L), and θ_{jn} parameters from the different analyses of GW190425 using a low spin prior limit ($\chi_{i,z} \in [-0.05, 0.05]$). The results include models TEOBResumS (blue) and SEOBNRv4T_surrogate (orange) done with RIFT and IMRPhenomPv2_NRTidal_ROQ (green) done with LALInference. The results are largely consistent across all waveforms.

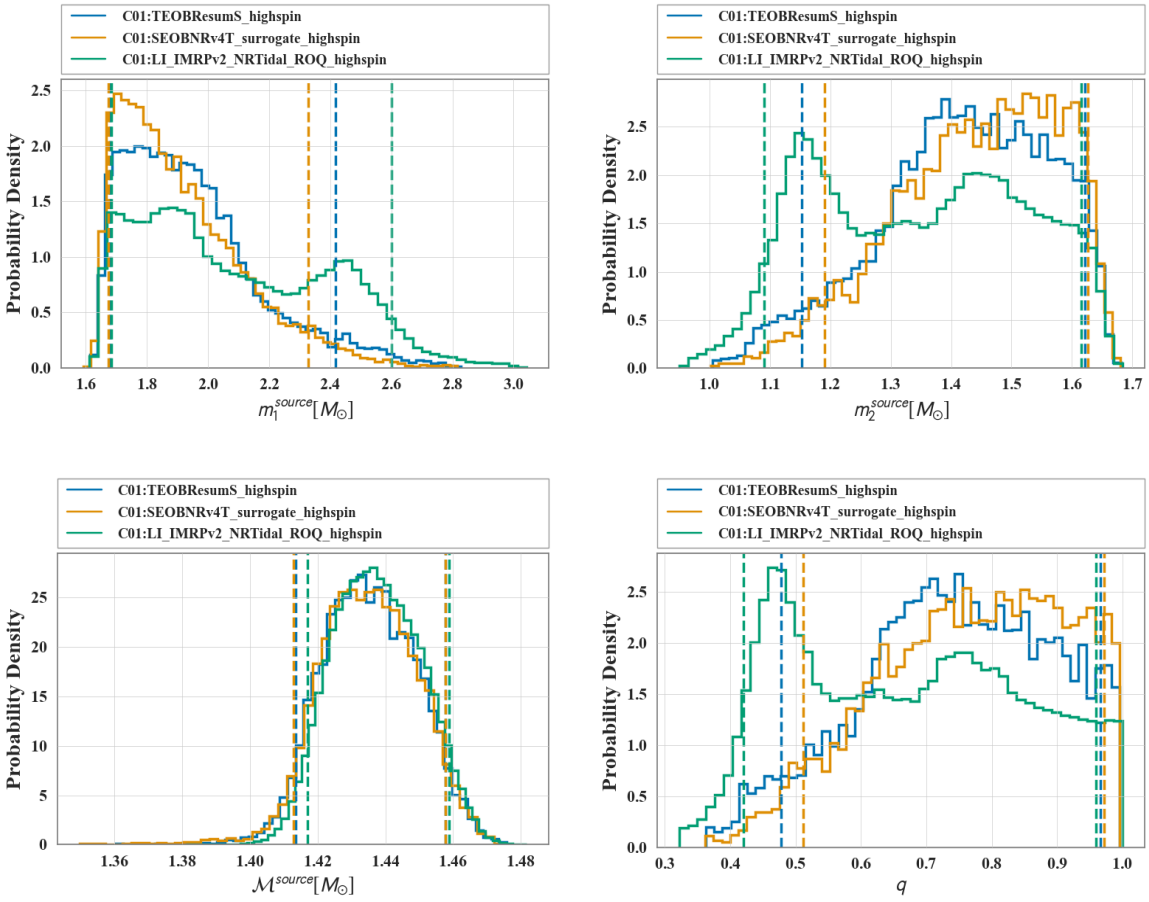

Figure 4.5: **GW190425 binary neutron star candidate with a high spin prior I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190425 using a highest possible spin limit. The results include models TEOBResumS (blue) and SEOBNRv4T_surrogate (orange) done with RIFT and IMRPhenomPv2_NRTidal_ROQ (green) done with LALInference. The EOB models peak closer to equal mass systems and omit a second peak at lower mass ratios compared to the IMRPhenomPv2_NRTidal_ROQ result.

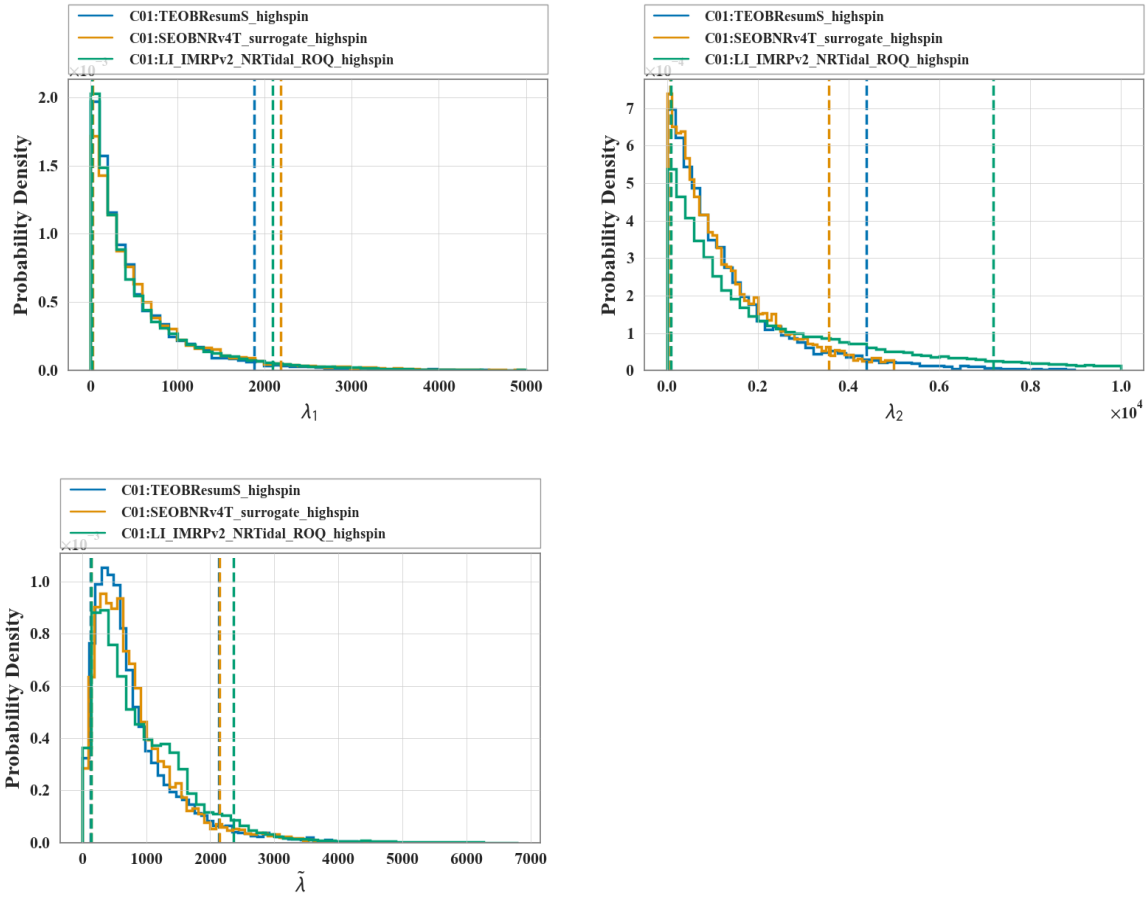

Figure 4.6: **GW190425 Potential Binary Neutron Star System with a high spin prior II:** This figure shows panels of the tidal parameters from the different analyses of GW190425 using a highest possible spin limit. The results include models TEOBRESUMS (blue) and SEOBNRV4T_SURROGATE (orange) done with RIFT and IMRPHENOMPV2_NRTIDAL_ROQ (green) done with LALInference. The EOB models omit a secondary peak at higher $\widetilde{\Lambda}$ compared to the IMRPHENOMPV2_NRTIDAL_ROQ result. Note SEOBNRV4T_SURROGATE can only generate waveforms $\lambda_i \leq 5000$.

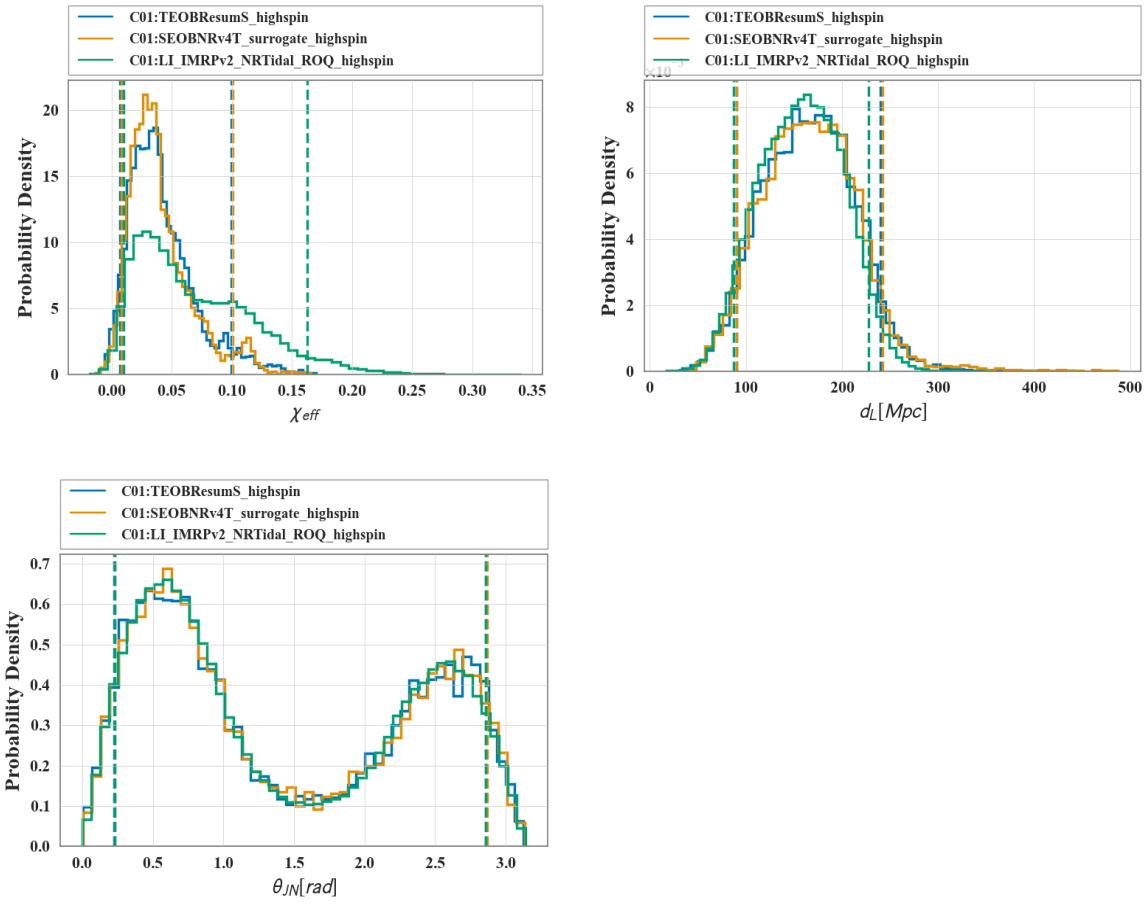

Figure 4.7: **GW190425 Binary Neutron Star System with a high spin prior III:** This figure shows panels of the 1D marginal distributions for the $\chi_{\rm eff}$, luminosity distance (d_L), and θ_{jn} parameters from the different analyses of GW190425 using a highest possible spin limit. The results include models TEOBResumS (blue) and SEOBNRv4T_surrogate (orange) done with RIFT and IMRPhenomPv2_NRTidal_ROQ (green) done with LALInference. The EOB models have a smaller secondary peak toward positive $\chi_{\rm eff}$; luminosity distance and θ_{jn} distributions are largely the same between waveforms.

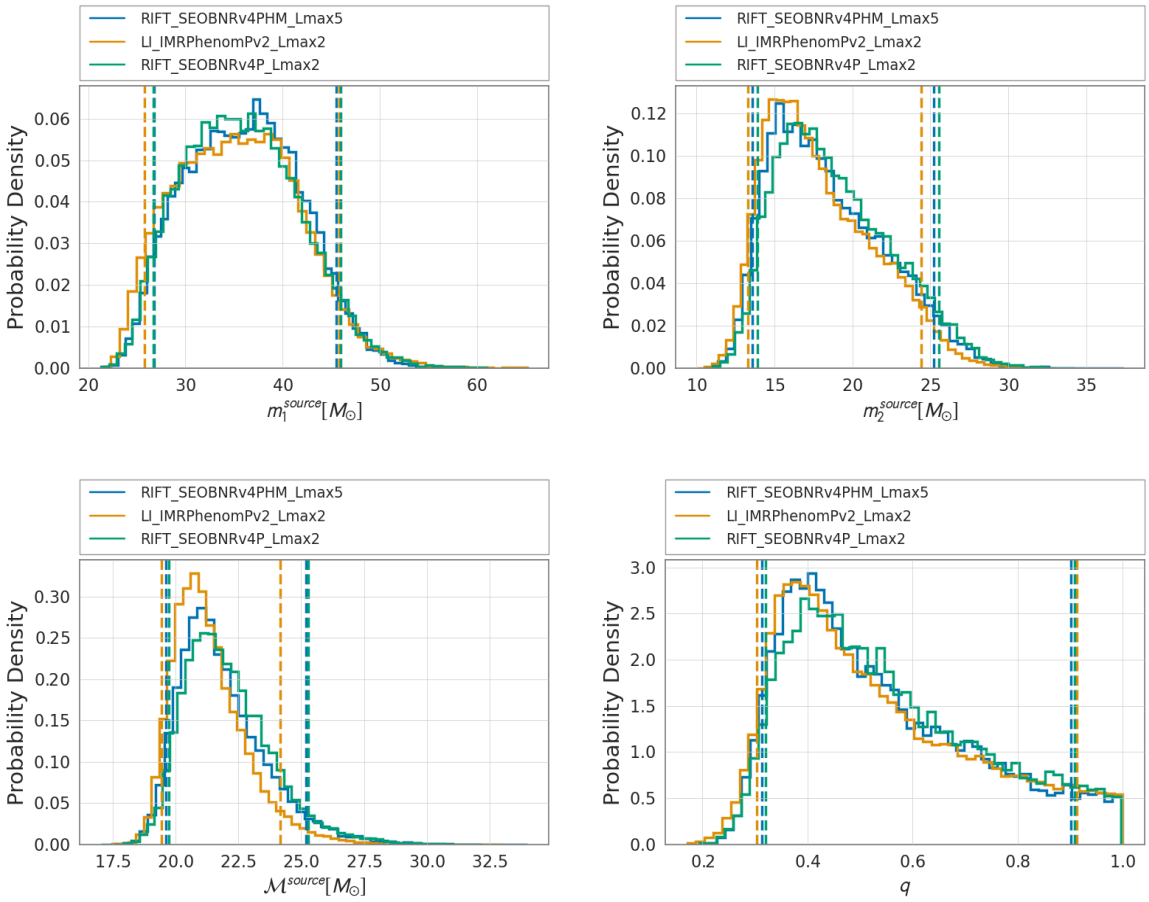

Figure 4.8: **GW190513 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190513. The results include models SEOBNRv4PHM (blue) and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference.

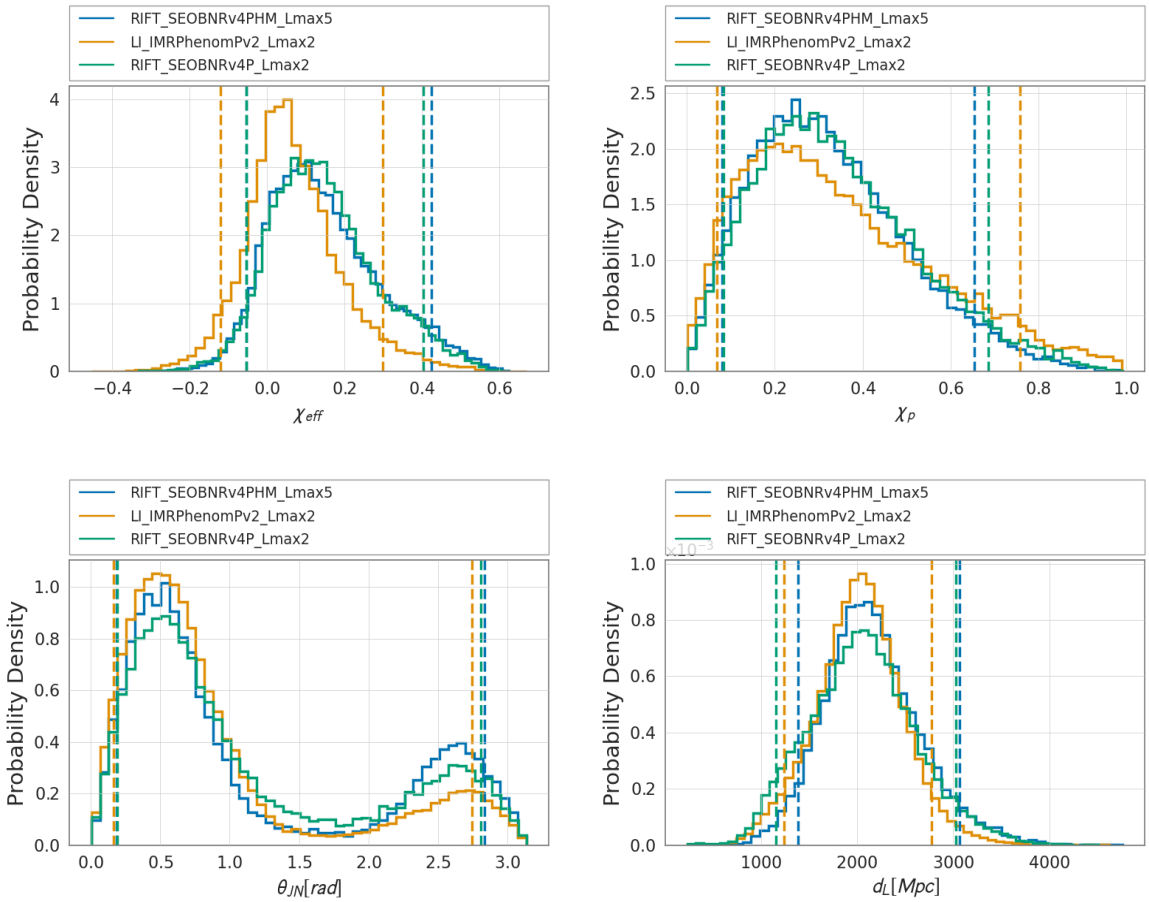

Figure 4.9: **GW190513 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190513. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between the two waveform families; marginal differences in posteriors that includes and omits higher order modes.

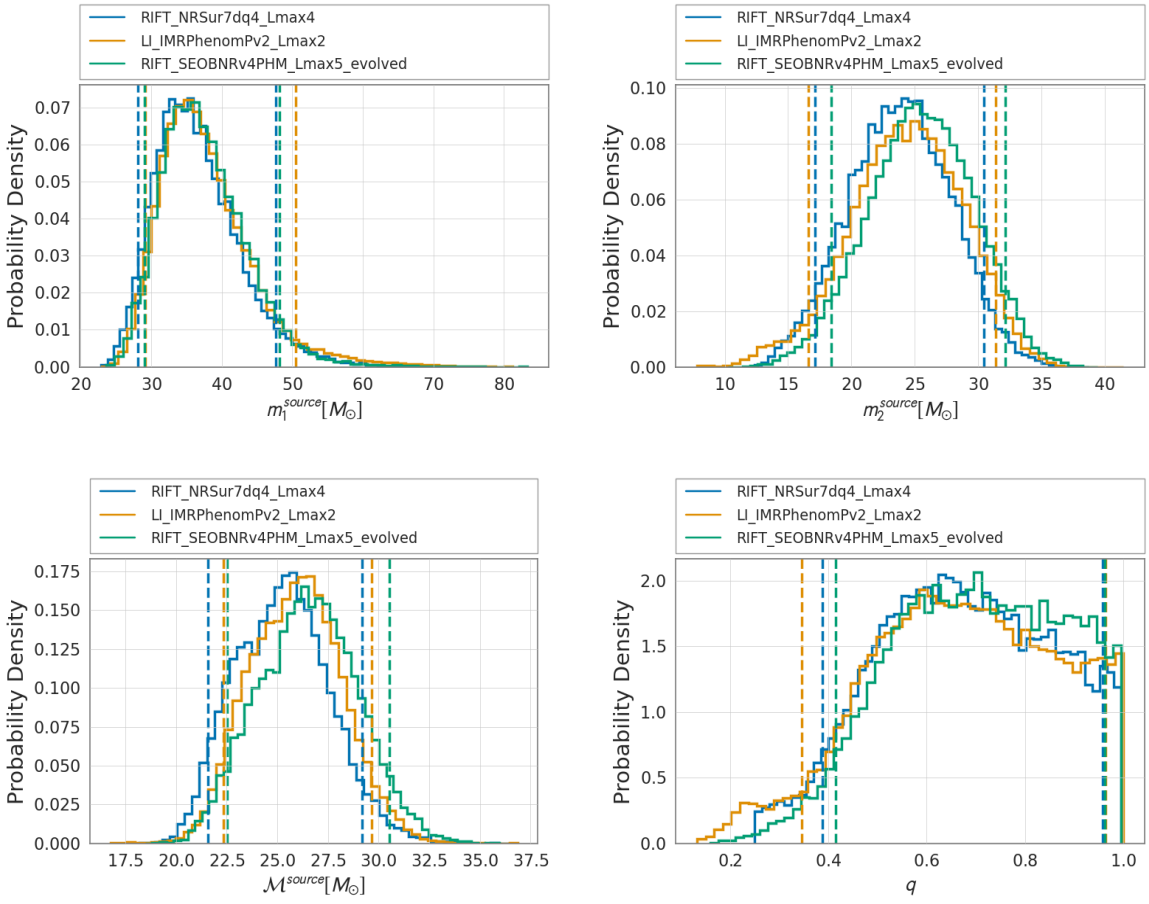

Figure 4.10: **GW190517 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190517. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms.

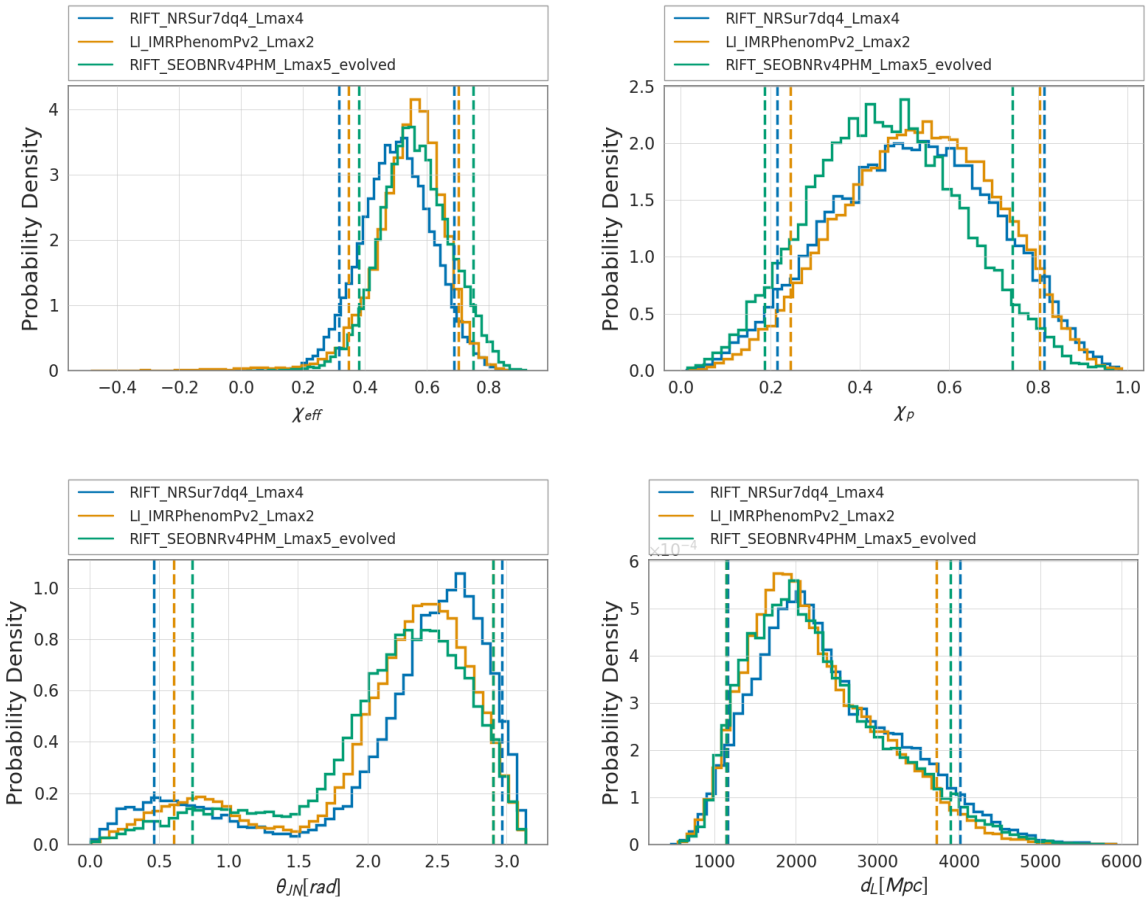

Figure 4.11: **GW190517 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190517. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms.

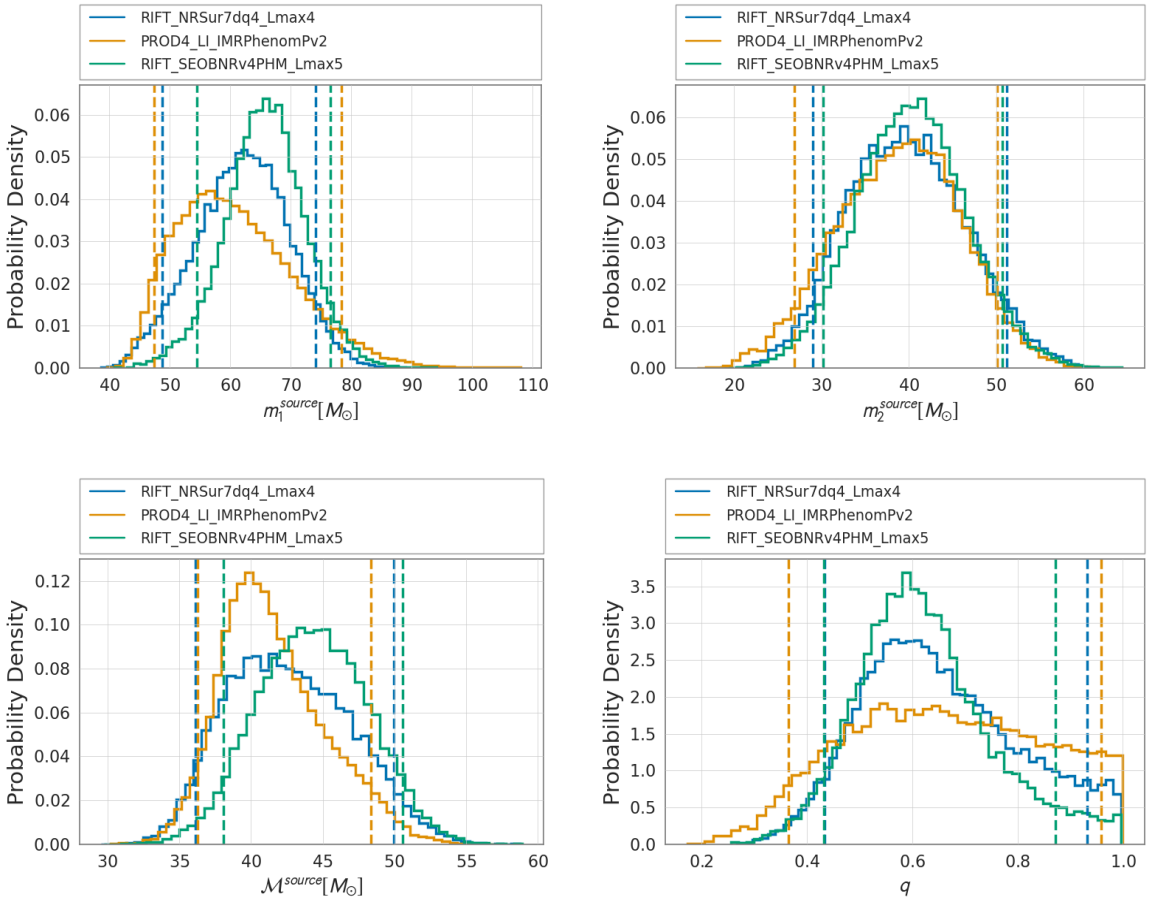

Figure 4.12: **GW190519 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190519. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Significant differences between different waveform families; significant differences when including and omit higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

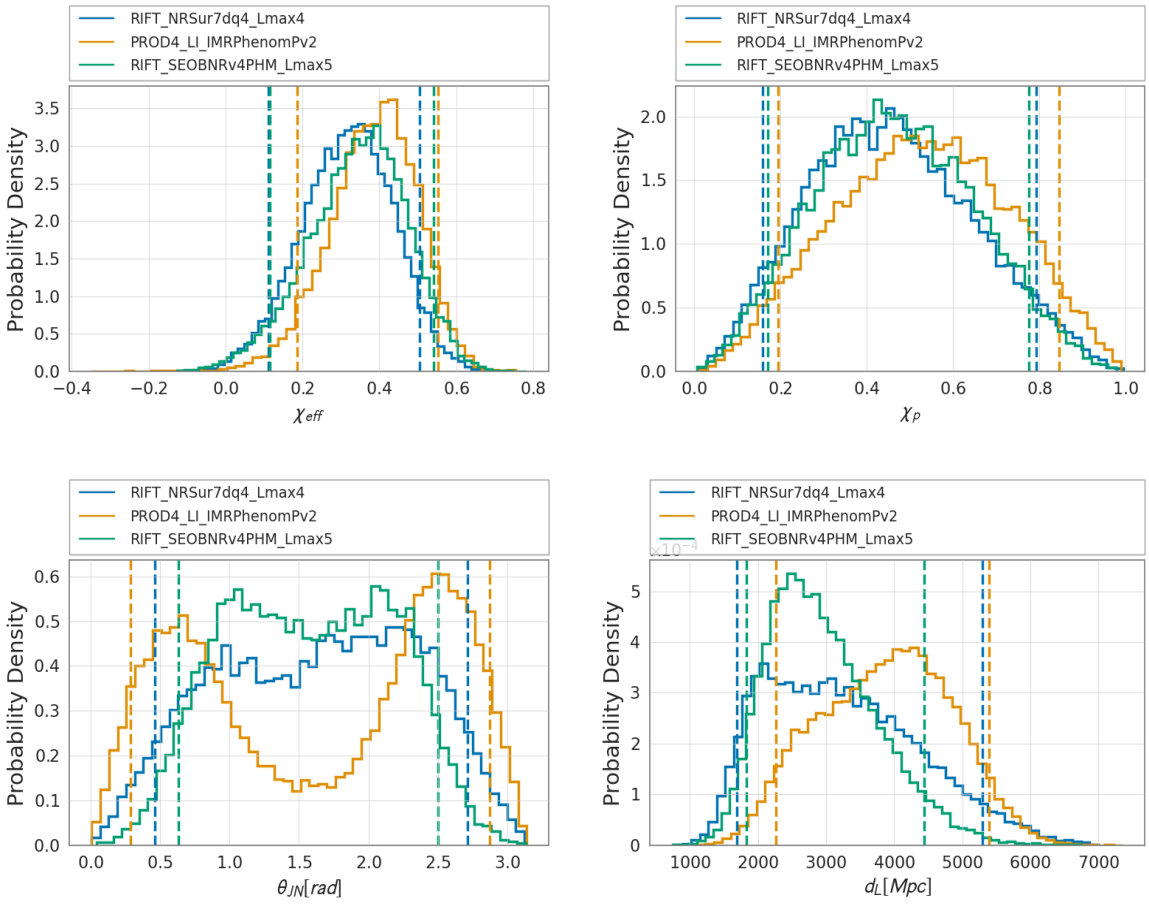

Figure 4.13: **GW190519 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190519. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Significant differences between different waveform families; significant differences when including and omit higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

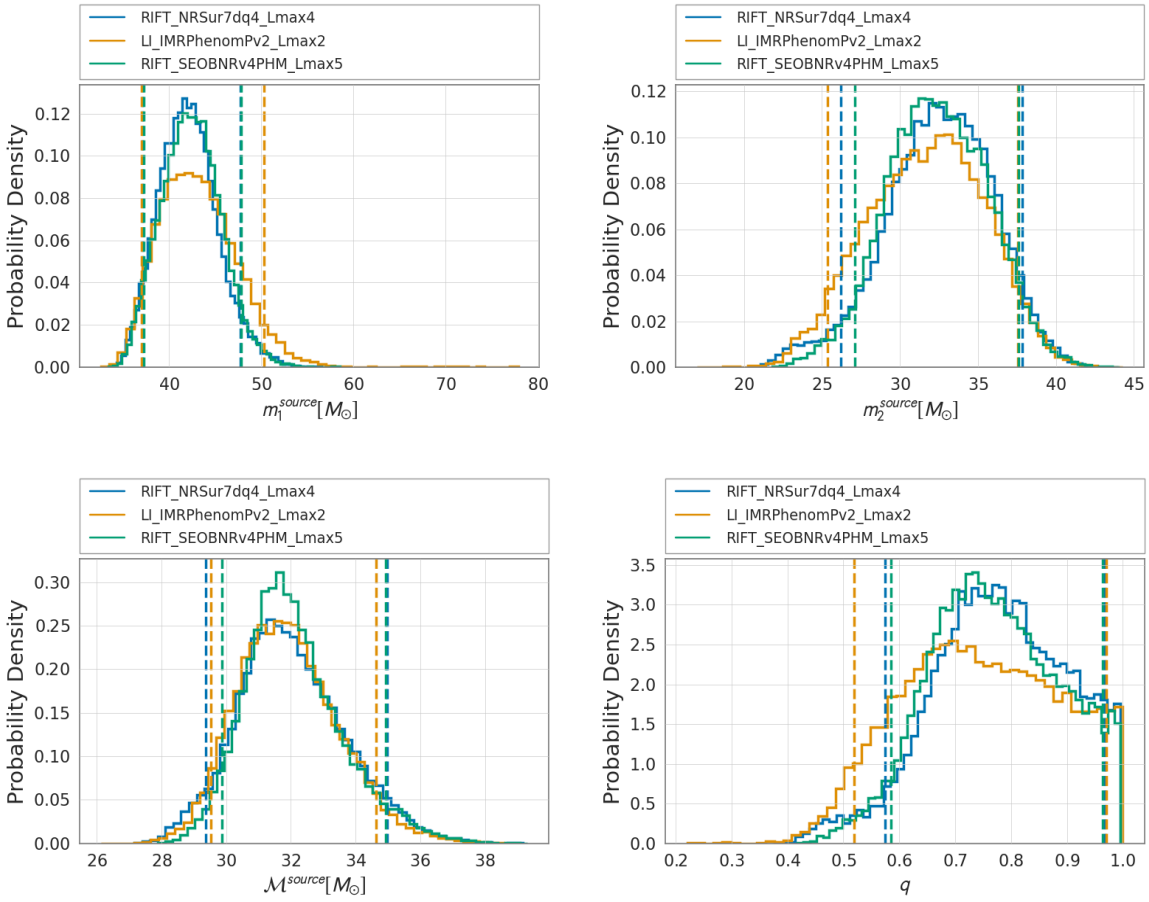

Figure 4.14: **GW190521r Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190521r. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between the different waveforms.

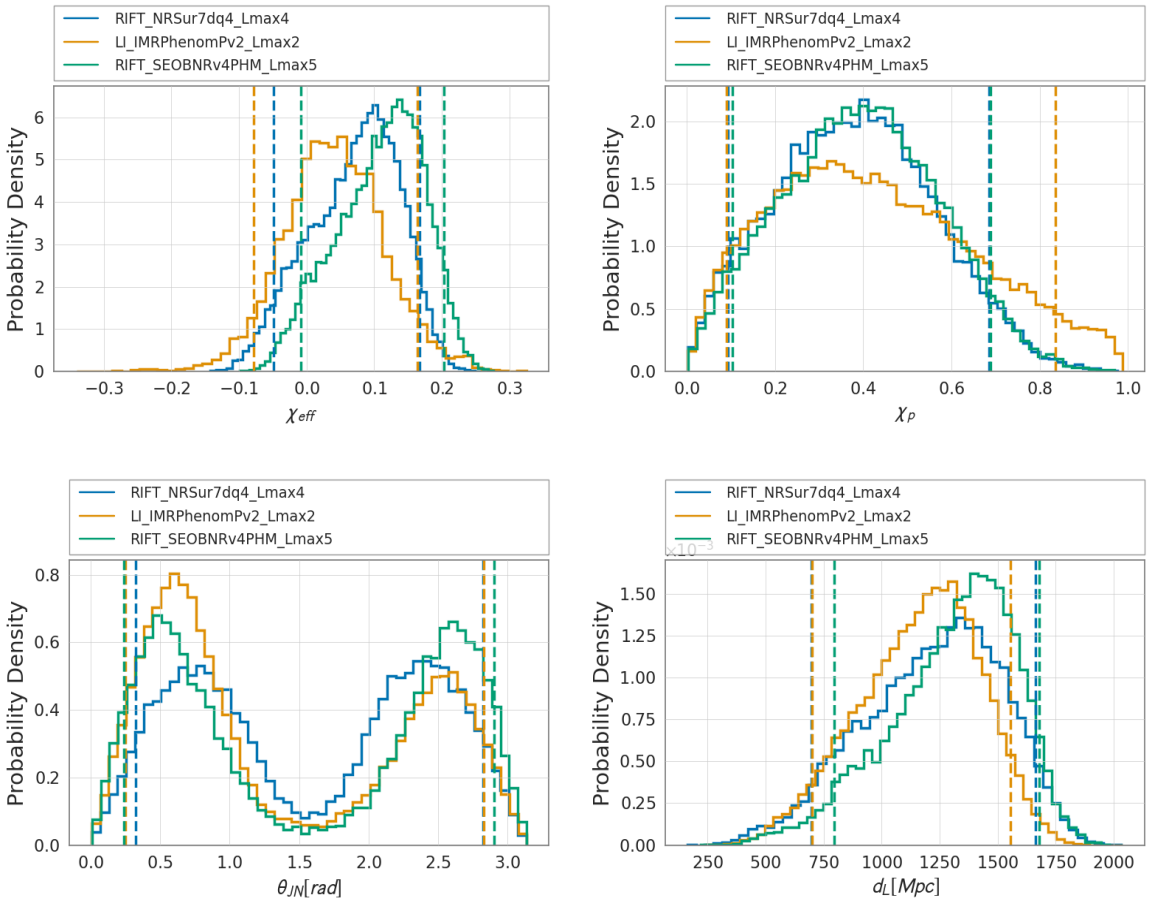

Figure 4.15: **GW190521r Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190521r. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between the different waveforms.

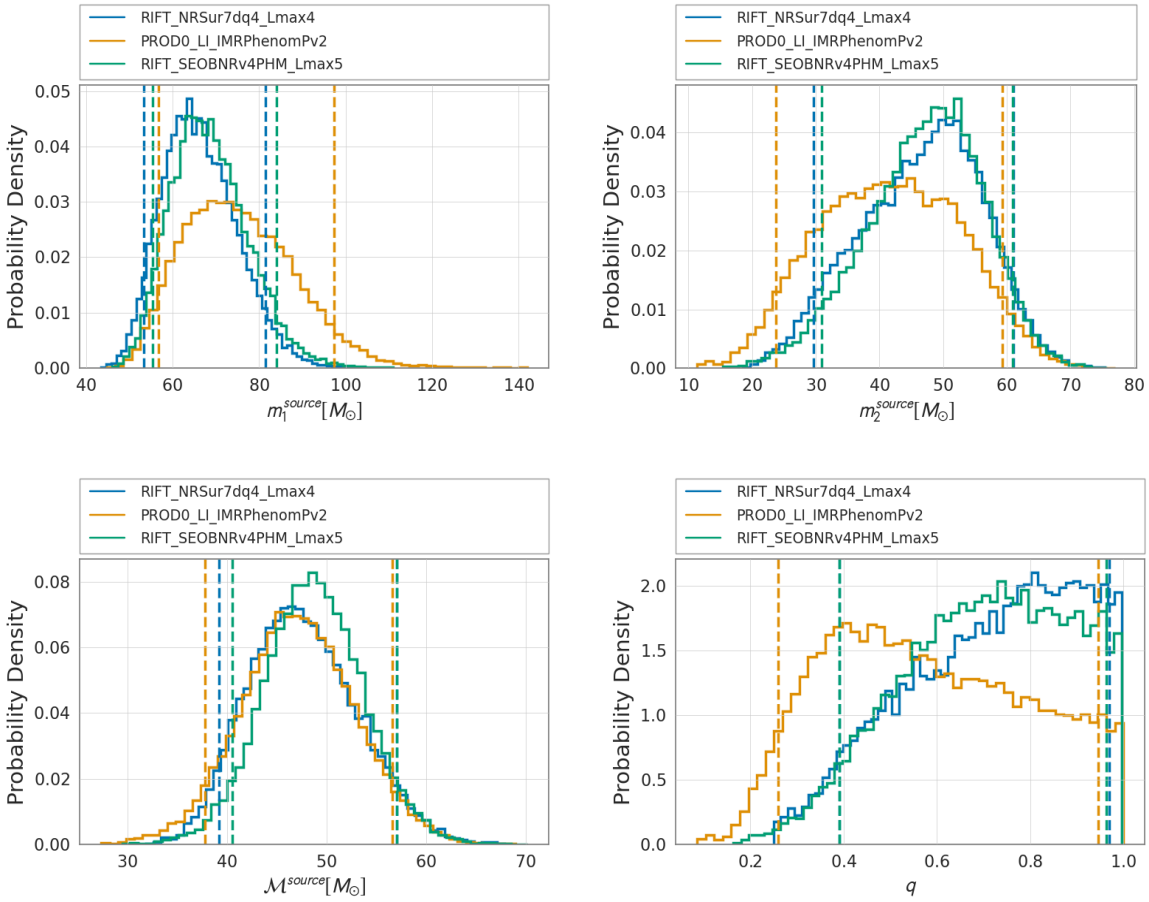

Figure 4.16: **GW190602 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190602. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Significant differences between different waveform families; significant differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

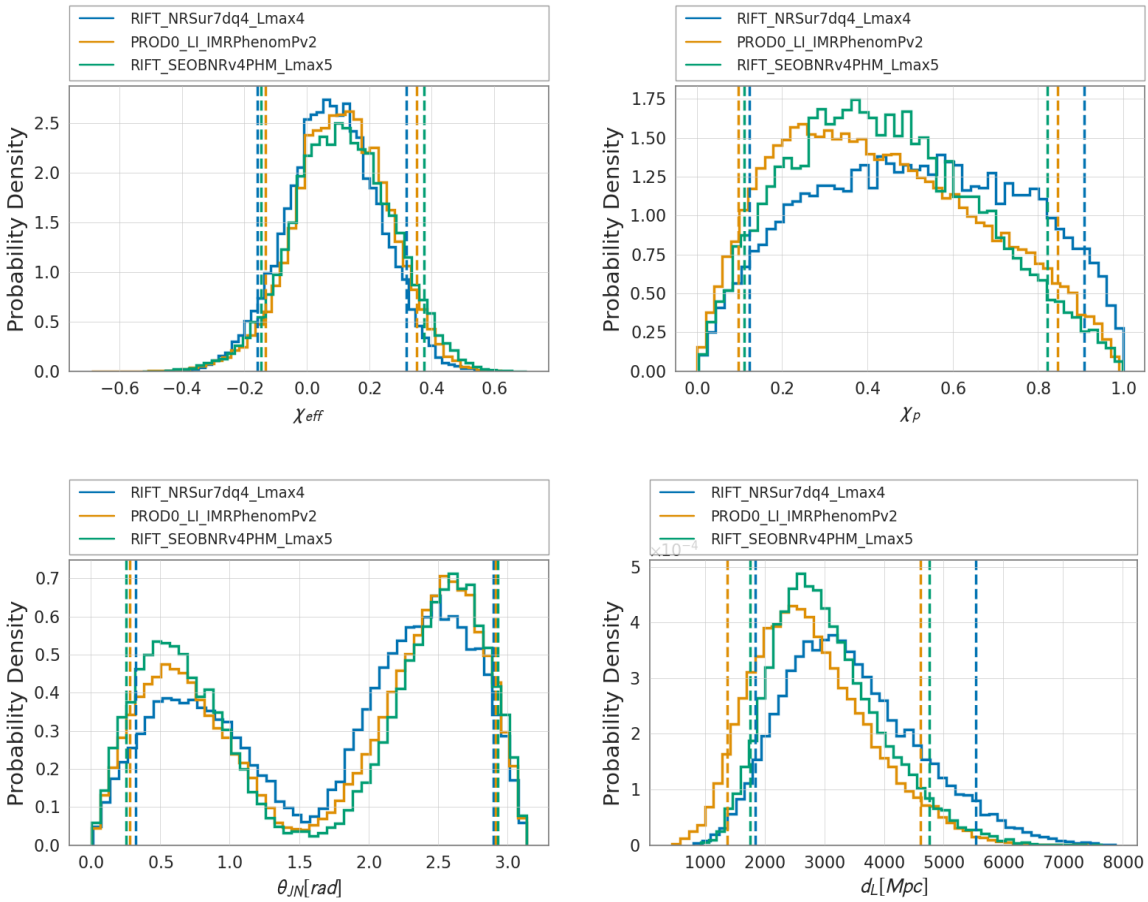

Figure 4.17: **GW190602 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190602. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Significant differences between different waveform families; significant differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

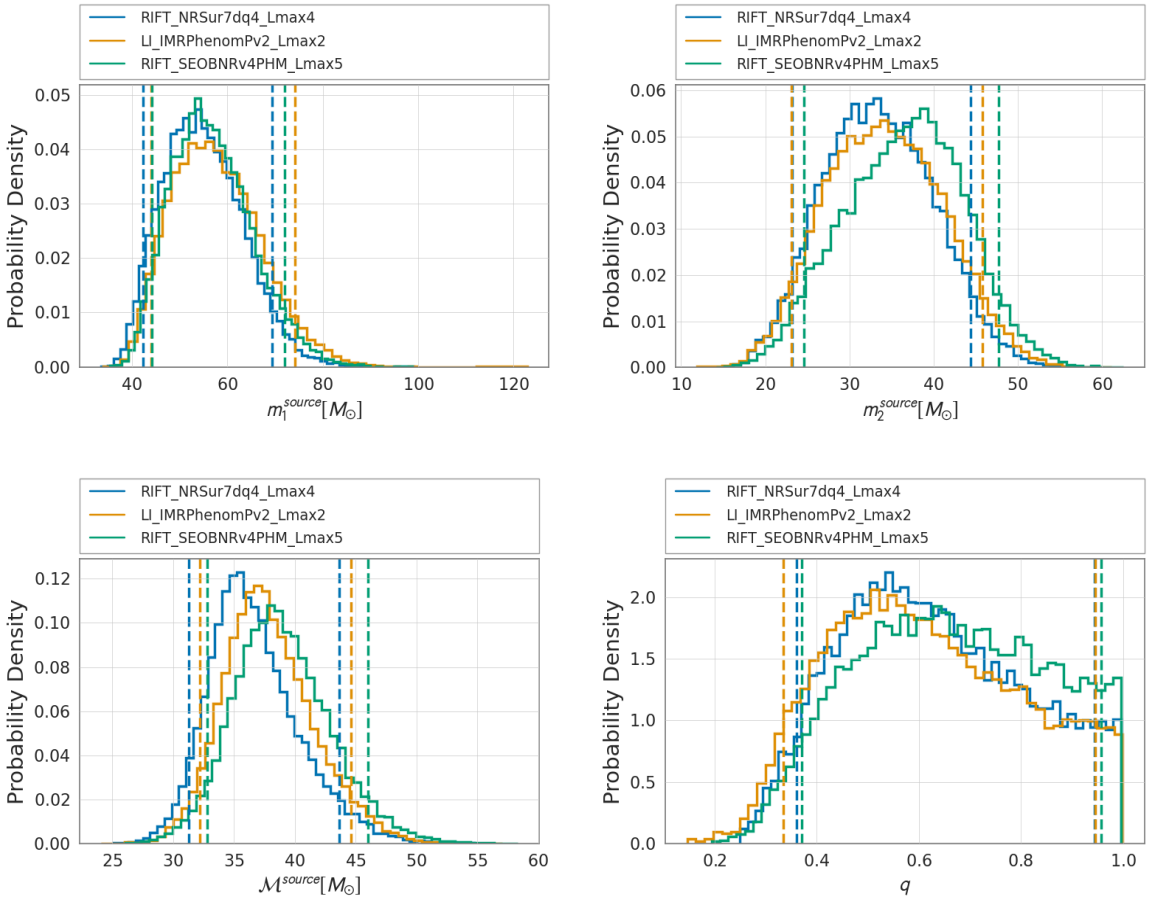

Figure 4.18: **GW190620 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190620. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveform families.

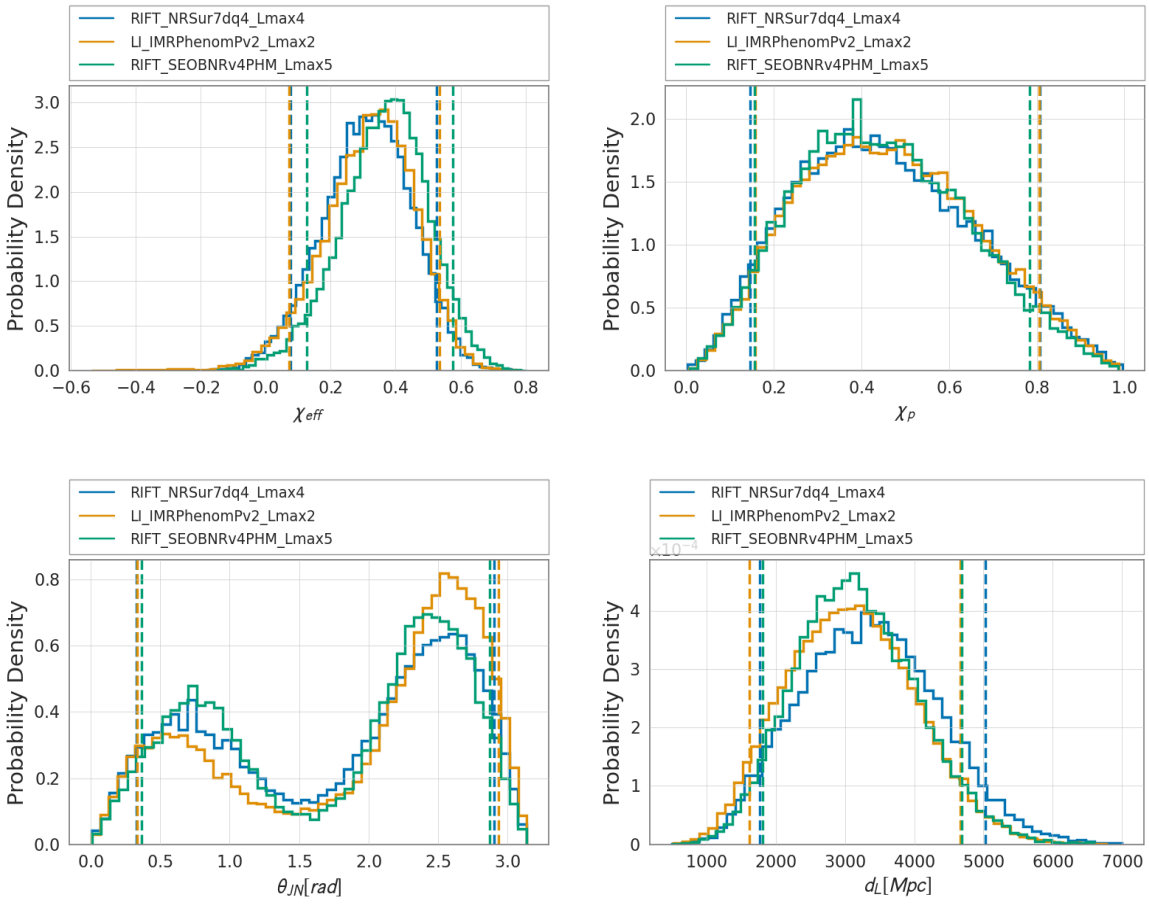

Figure 4.19: **GW190620 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190620. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveform families.

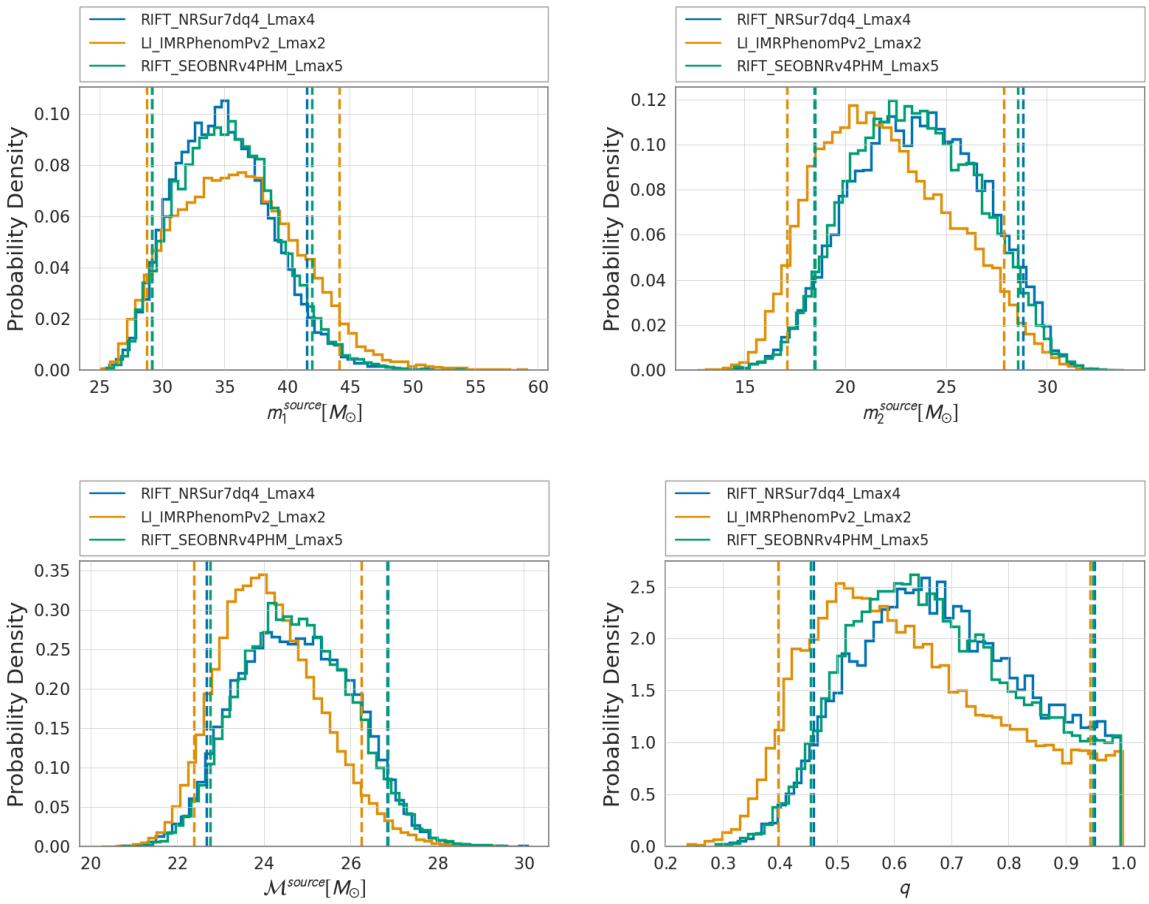

Figure 4.20: **GW190630 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190630. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences when including and omitting higher order modes.

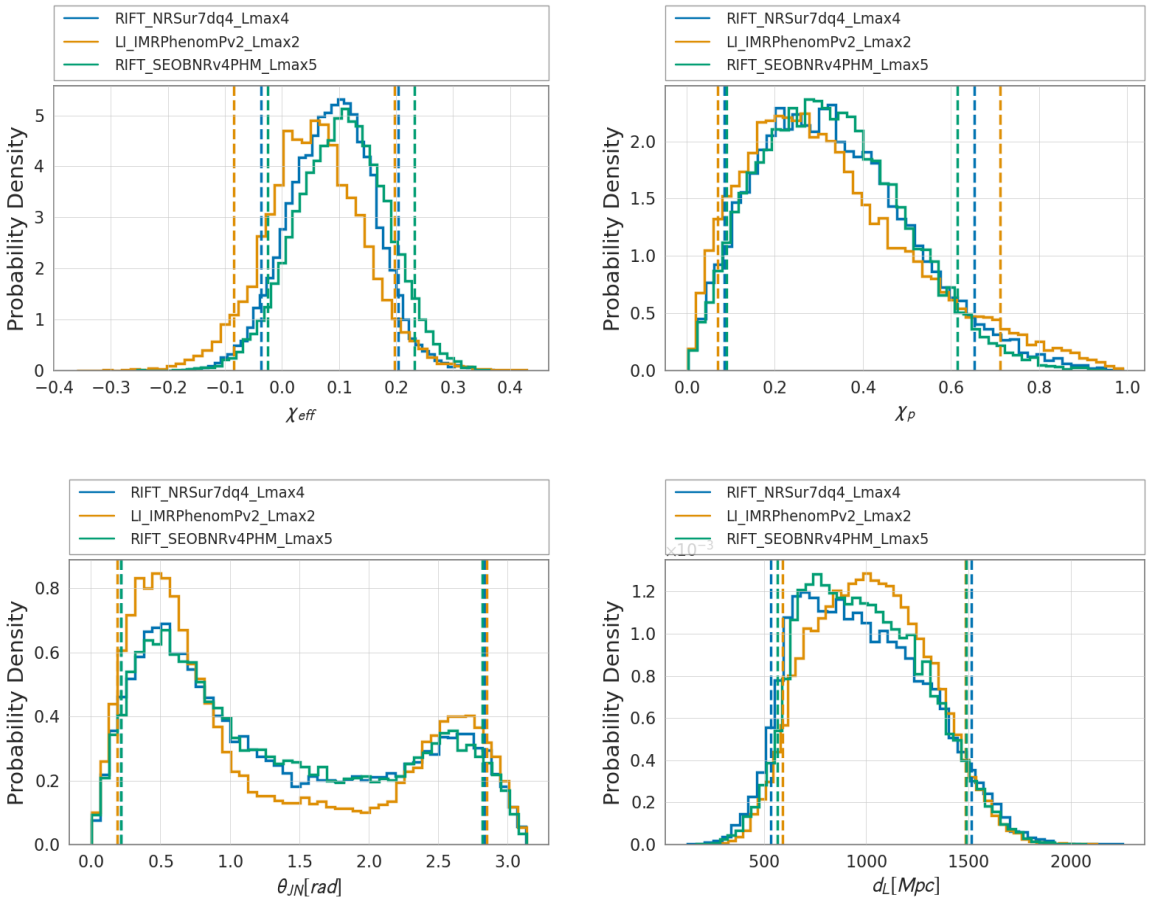

Figure 4.21: **GW190630 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190630. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveform families.

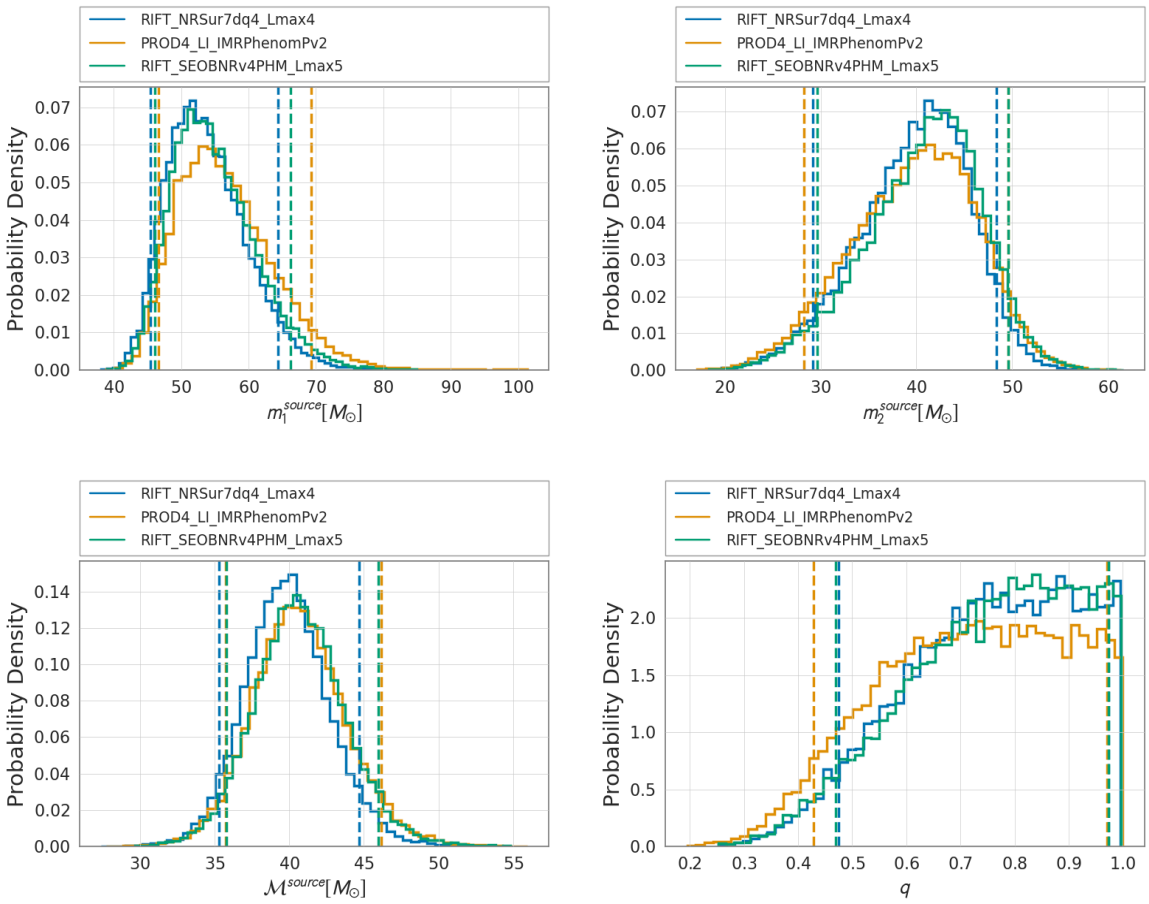

Figure 4.22: **GW190701 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190701. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences when including and omitting higher order modes.

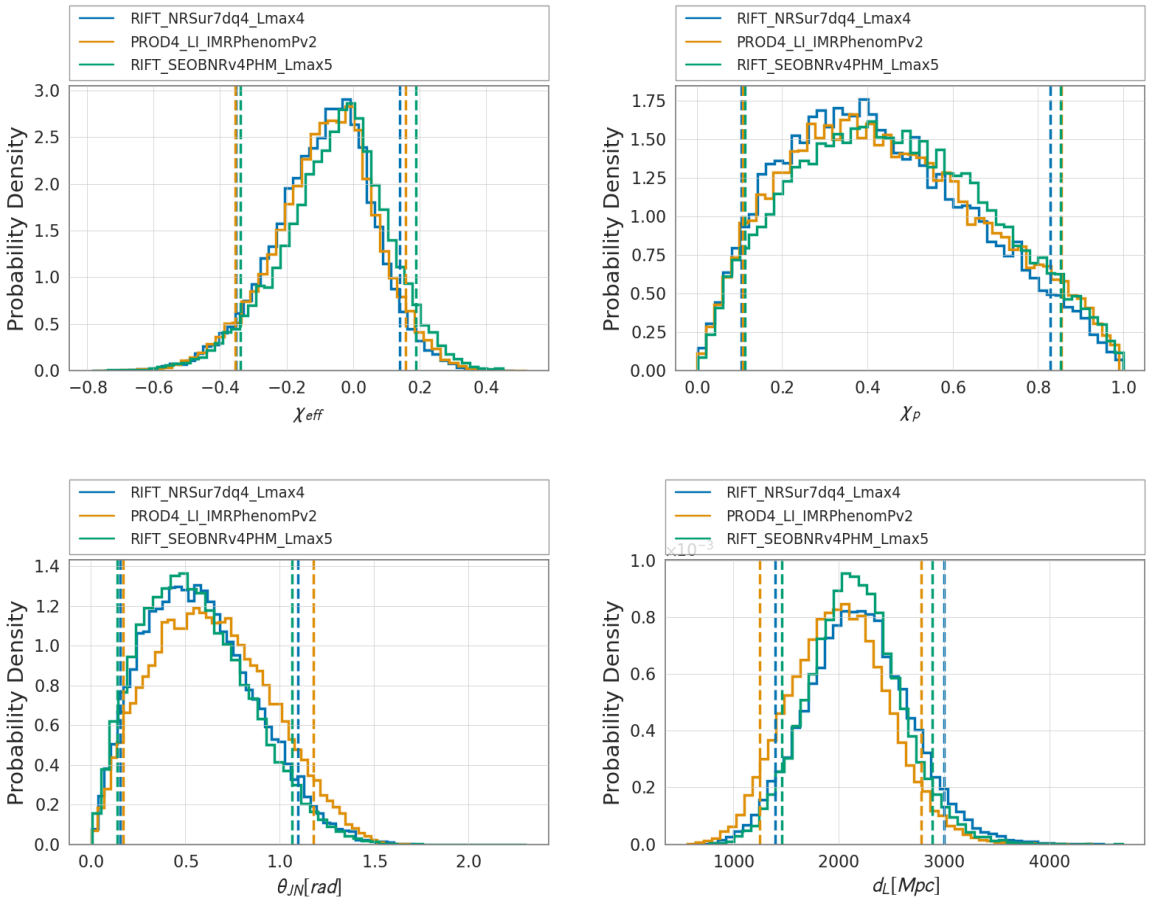

Figure 4.23: **GW190701 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190701. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences when including and omitting higher order modes.

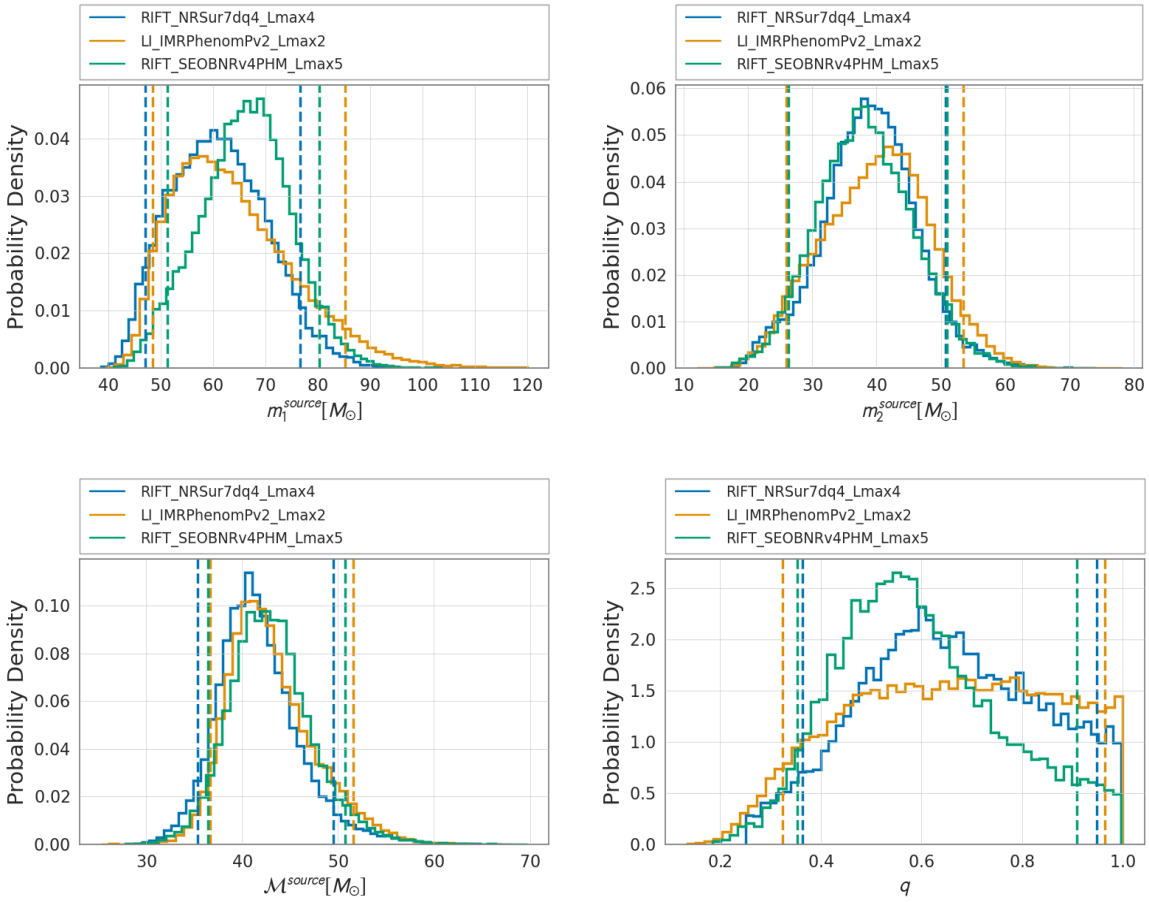

Figure 4.24: **GW190706 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190706. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between different waveforms; noticeable differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

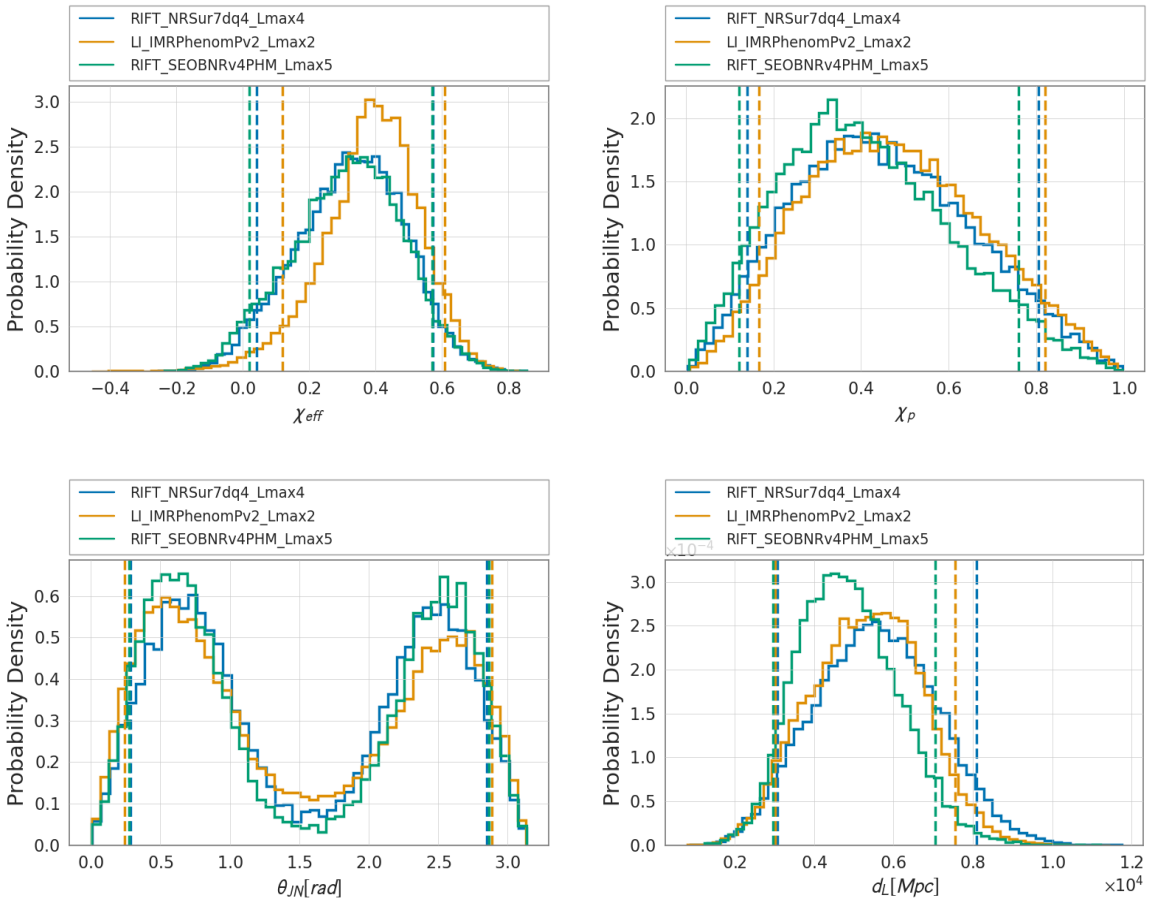

Figure 4.25: **GW190706 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190706. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between different waveforms; noticeable differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

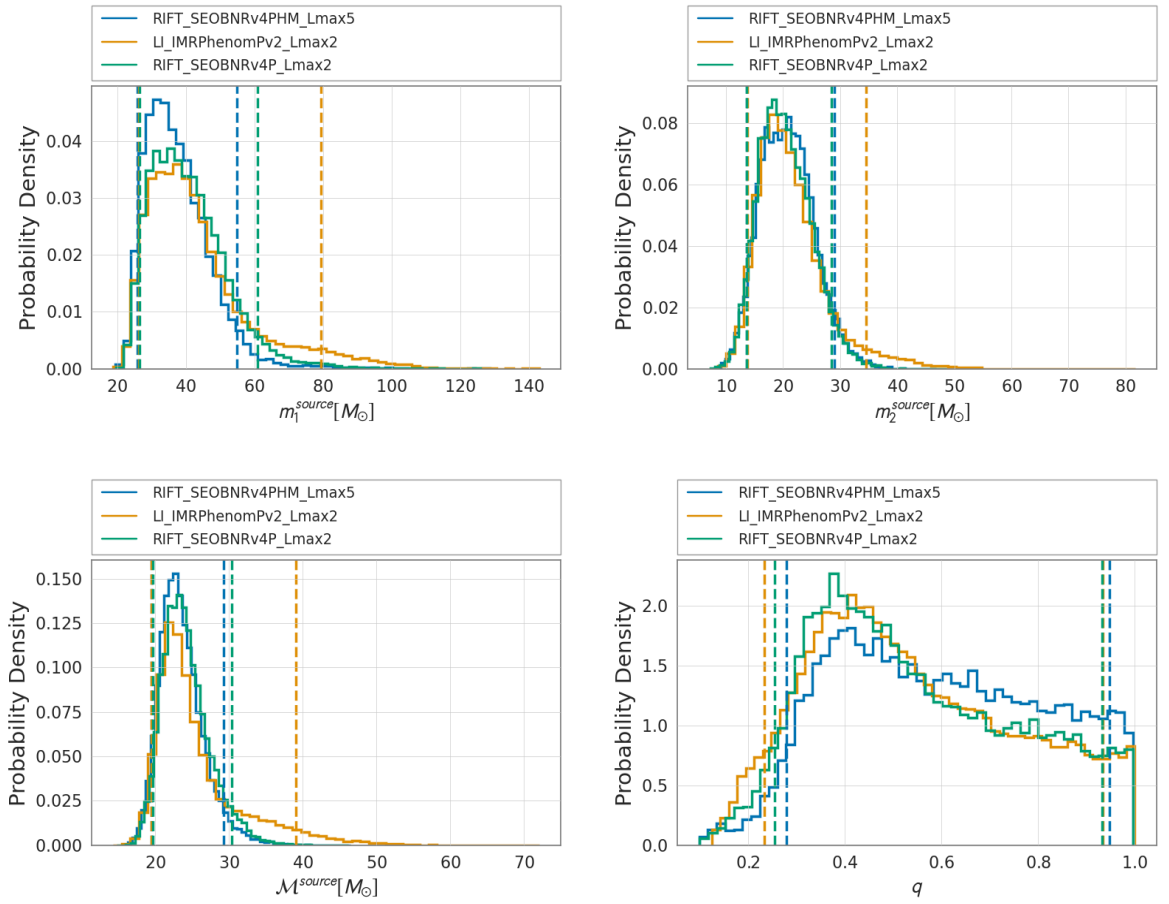

Figure 4.26: **GW190719 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190719. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; marginal differences when including and omitting higher order modes.

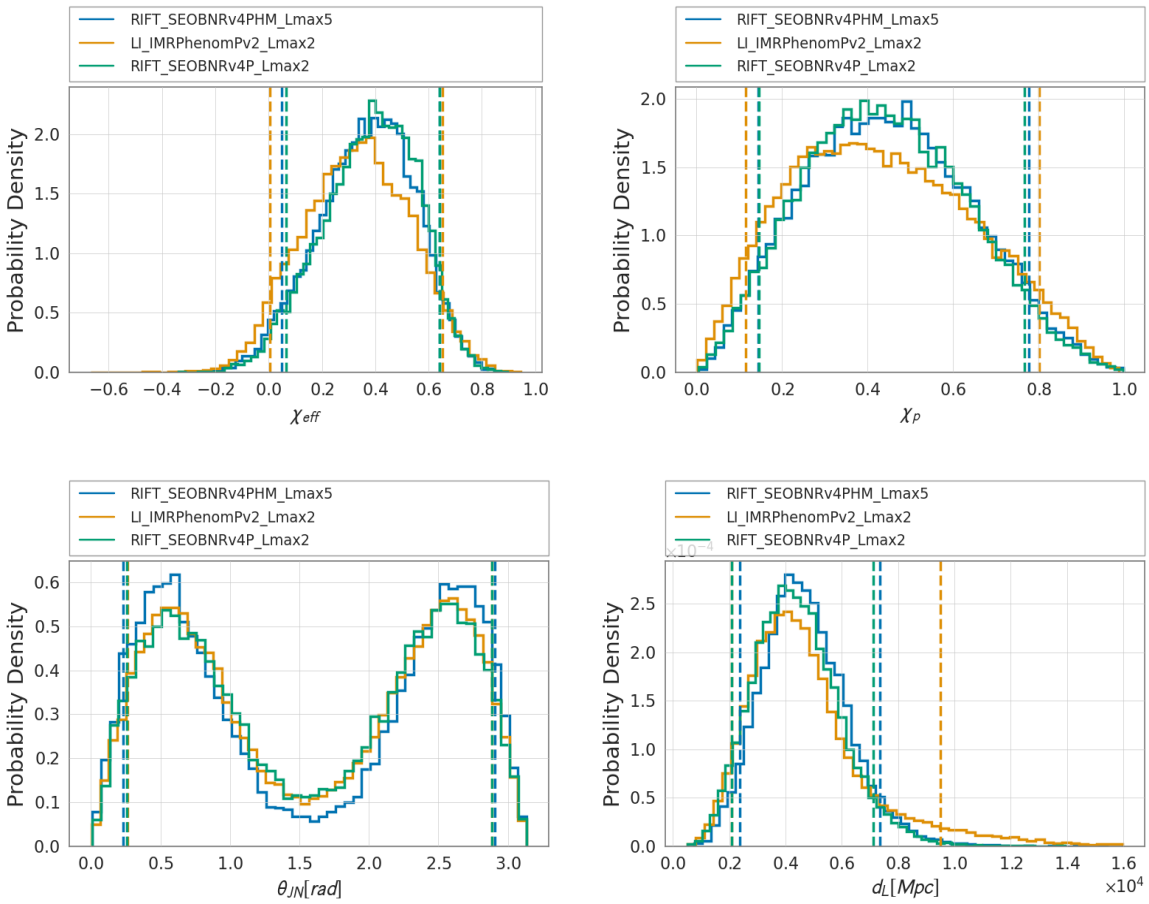

Figure 4.27: **GW190719 Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190719. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; marginal differences when including and omitting higher order modes.

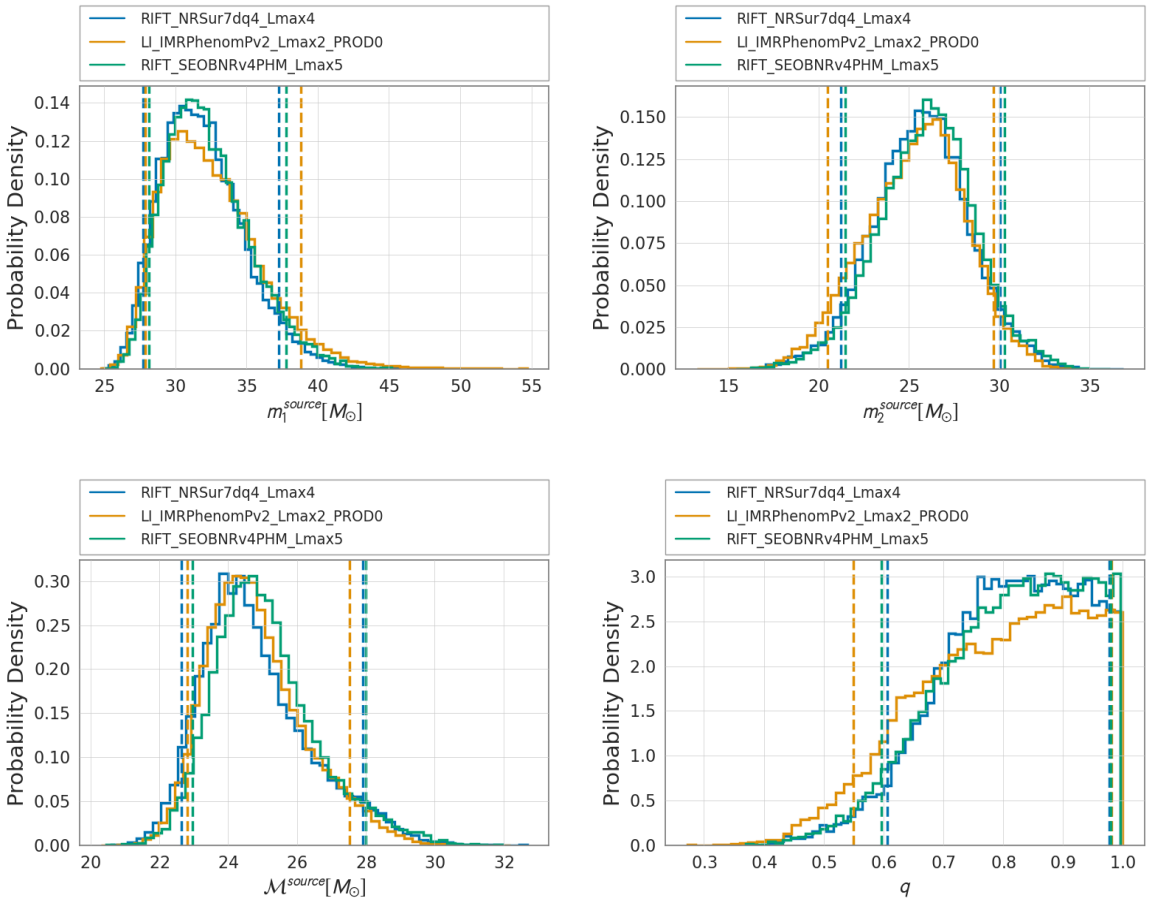

Figure 4.28: **GW190828j Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190828j. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; marginal differences when including and omitting higher order modes.

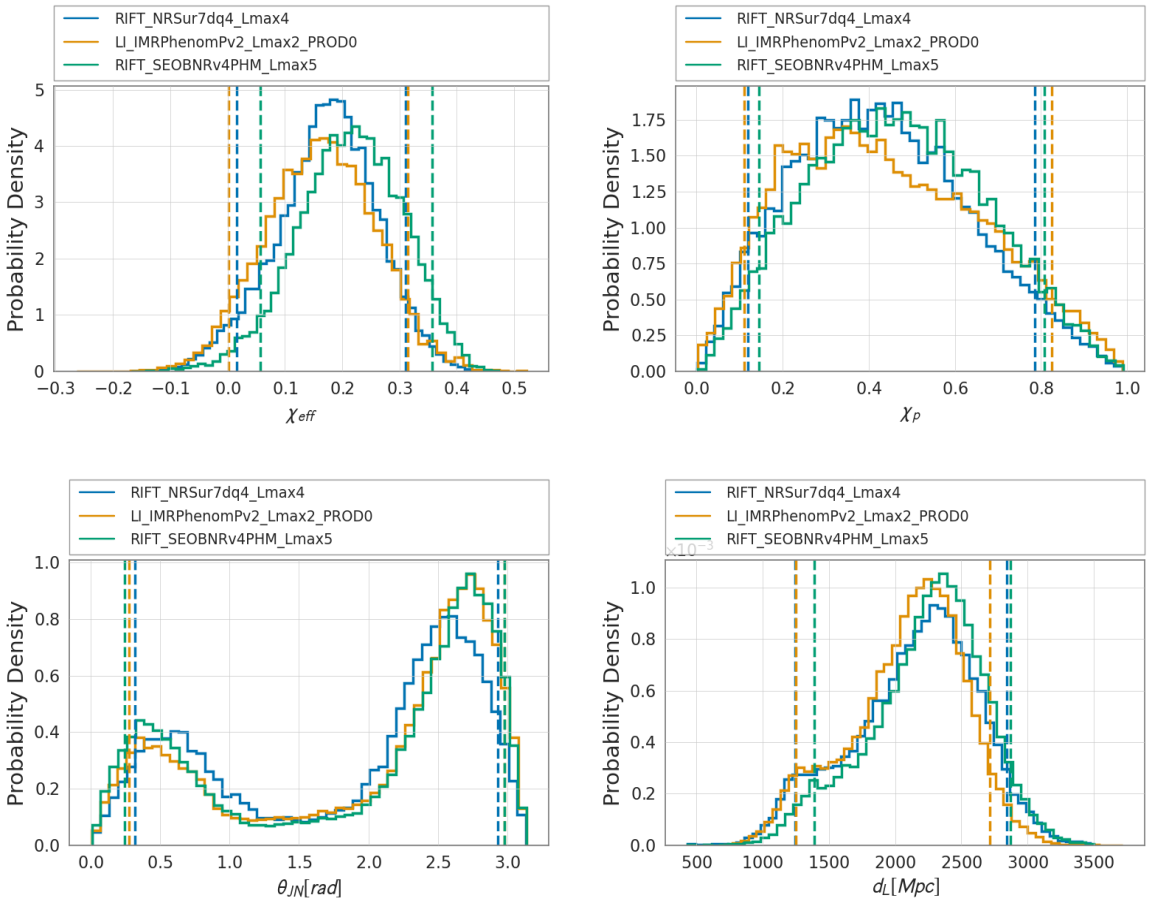

Figure 4.29: **GW190828j Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190828j. The results include models NRSur7dq4 (blue) with all available modes and SEOBNRv4PHM (green) with all available modes done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; marginal differences when including and omitting higher order modes.

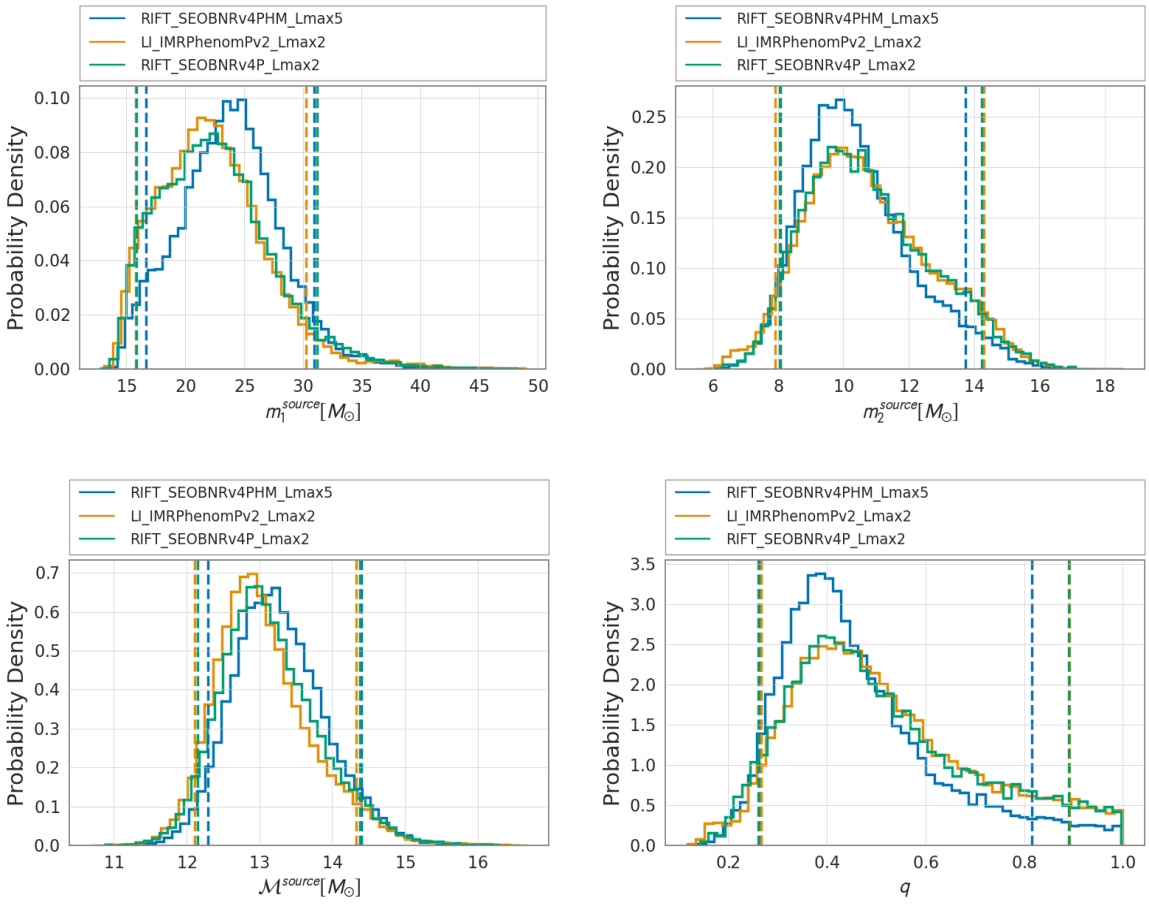

Figure 4.30: **GW190828ℓ Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190828ℓ. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; noticeable differences when including and omitting higher order modes.

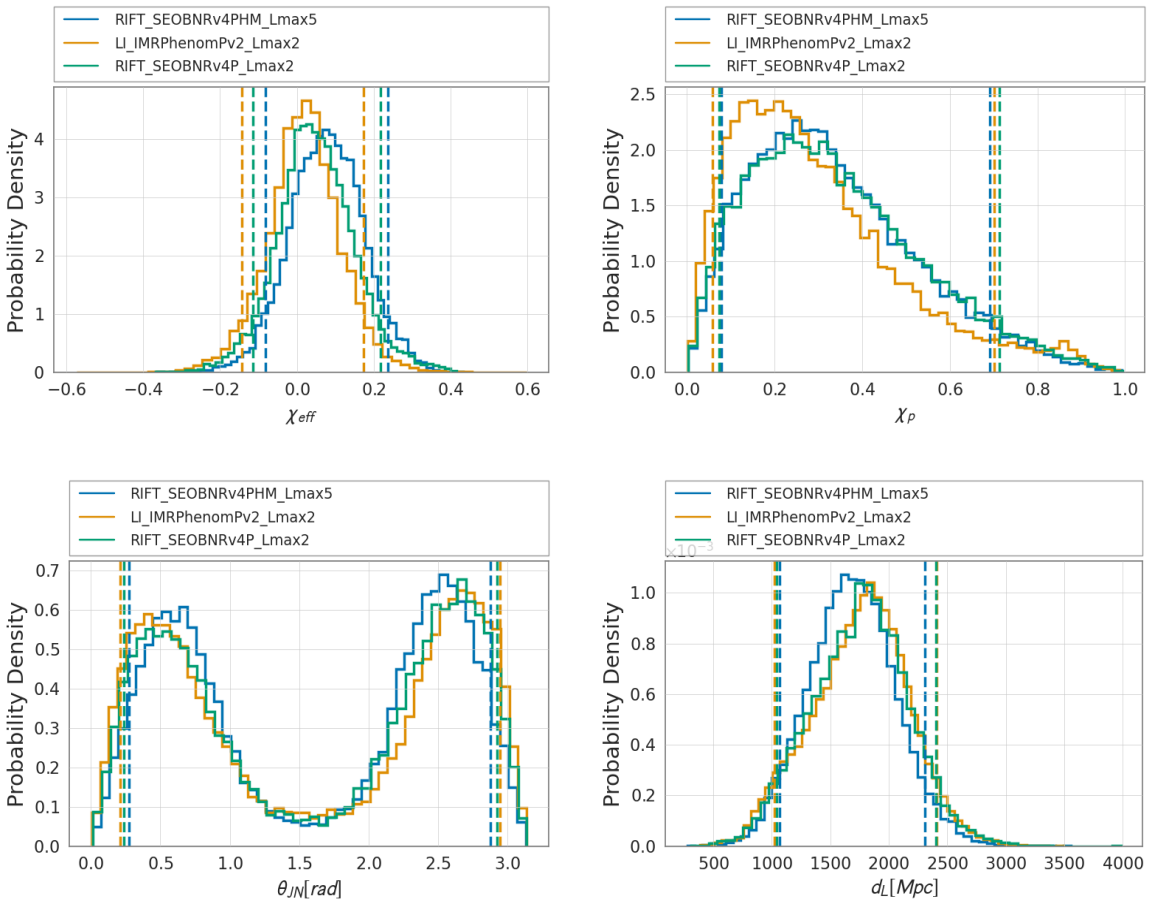

Figure 4.31: **GW190828ℓ Binary Black Hole System II:** This figure shows panels of the spins, θ_{jn}, and d_L parameters from the different analyses of GW190828ℓ. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between different waveforms; noticeable differences when including and omitting higher order modes.

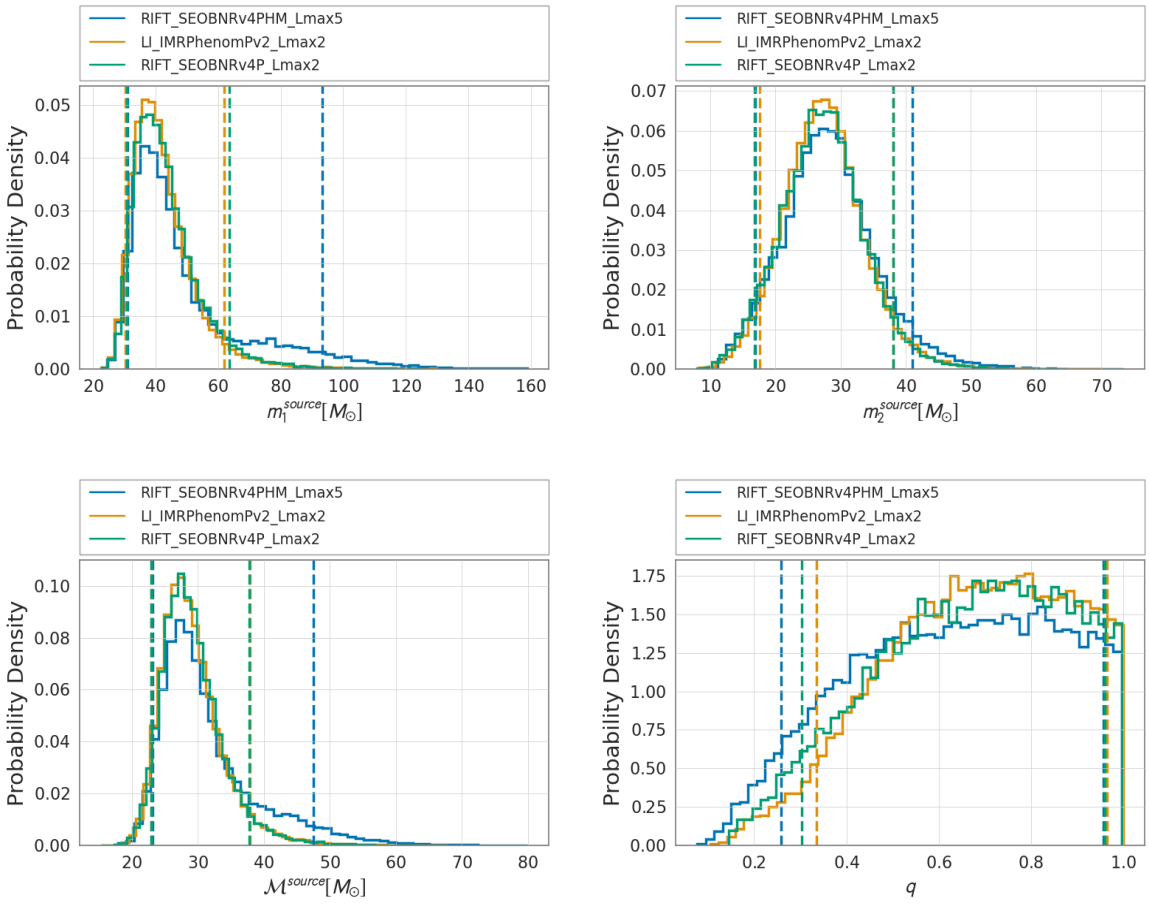

Figure 4.32: **GW190909 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190909. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; noticeable differences when including and omitting higher order modes.

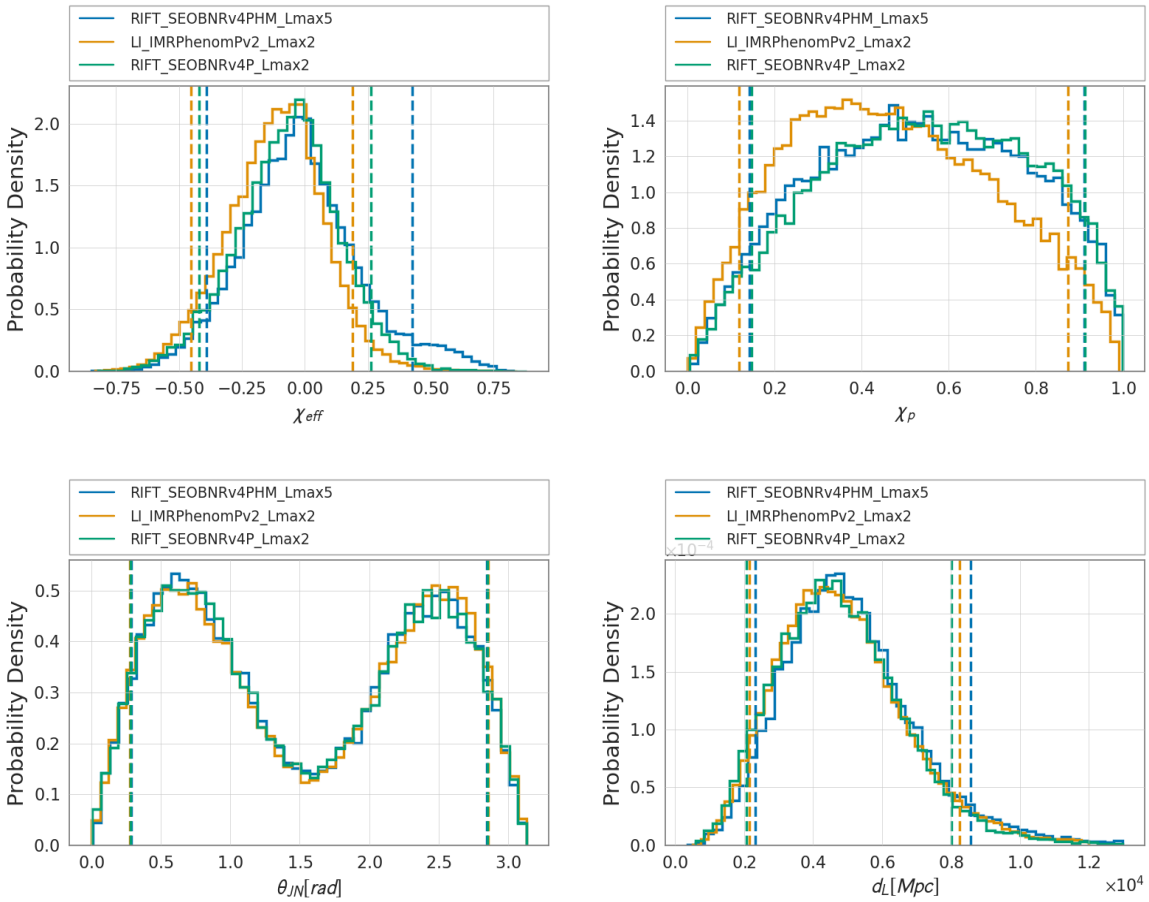

Figure 4.33: **GW190909 Binary Black Hole System II:** This figure shows panels of the 1D marginal distributions for the spins, θ_{jn}, and d_L parameters from the different analyses of GW190909. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between different waveforms; noticeable differences when including and omitting higher order modes.

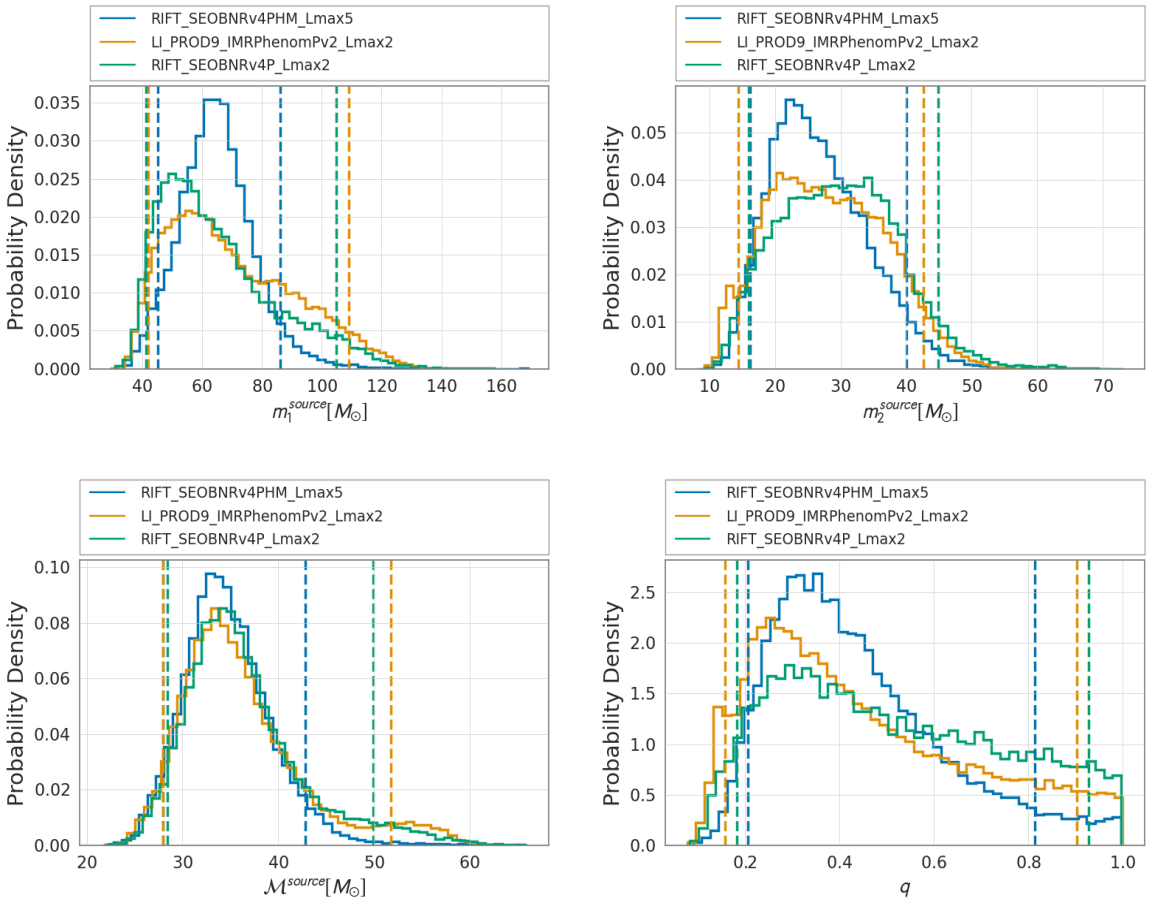

Figure 4.34: **GW190929 Binary Black Hole System I:** This figure shows panels of the 1D marginal distributions for the source frame mass parameters from the different analyses of GW190929. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Marginal differences between different waveforms; noticeable differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

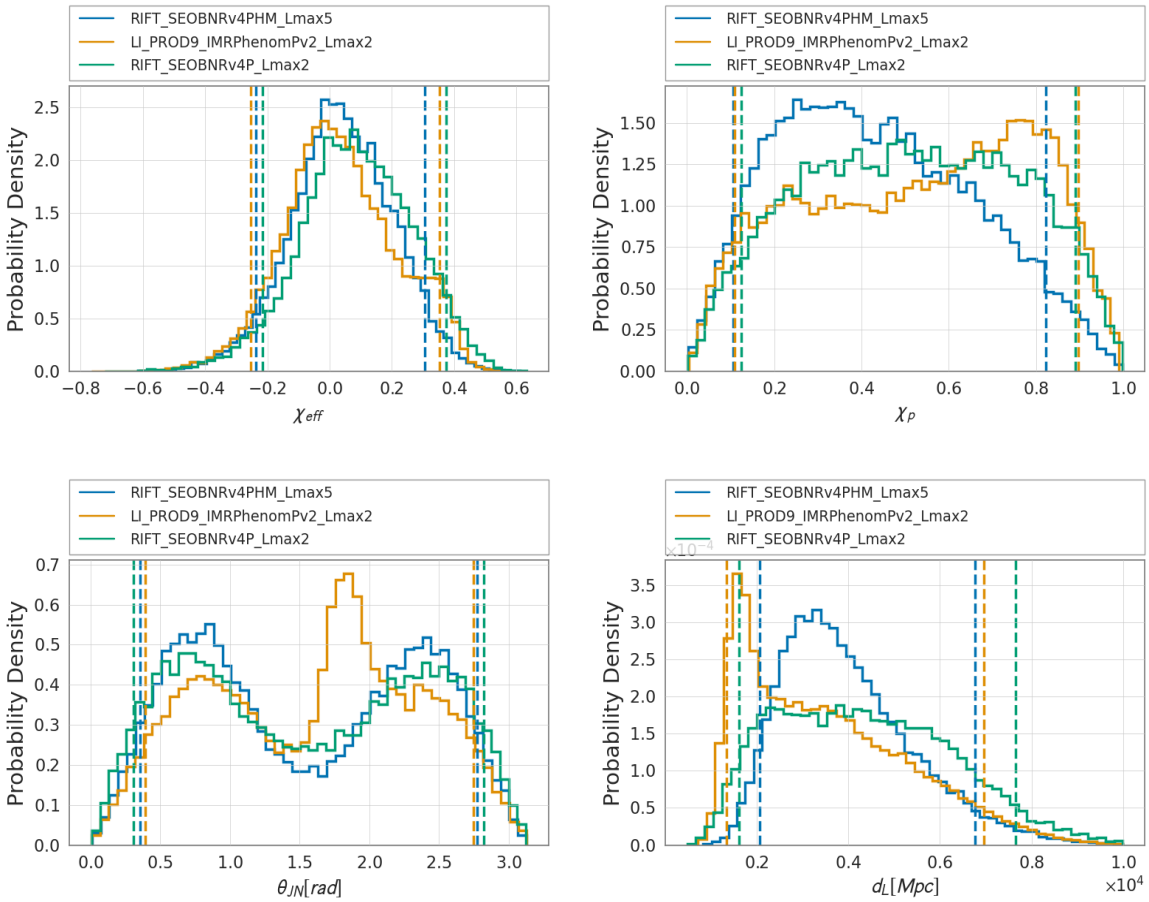

Figure 4.35: **GW190929 Binary Black Hole System II:** This figure shows panels of the 1D marginal distributions for the spins, θ_{jn}, and d_L parameters from the different analyses of GW190929. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. Noticeable differences between different waveforms; noticeable differences when including and omitting higher order modes. See Figure 4.36 for a more detailed comparison of the HOMs.

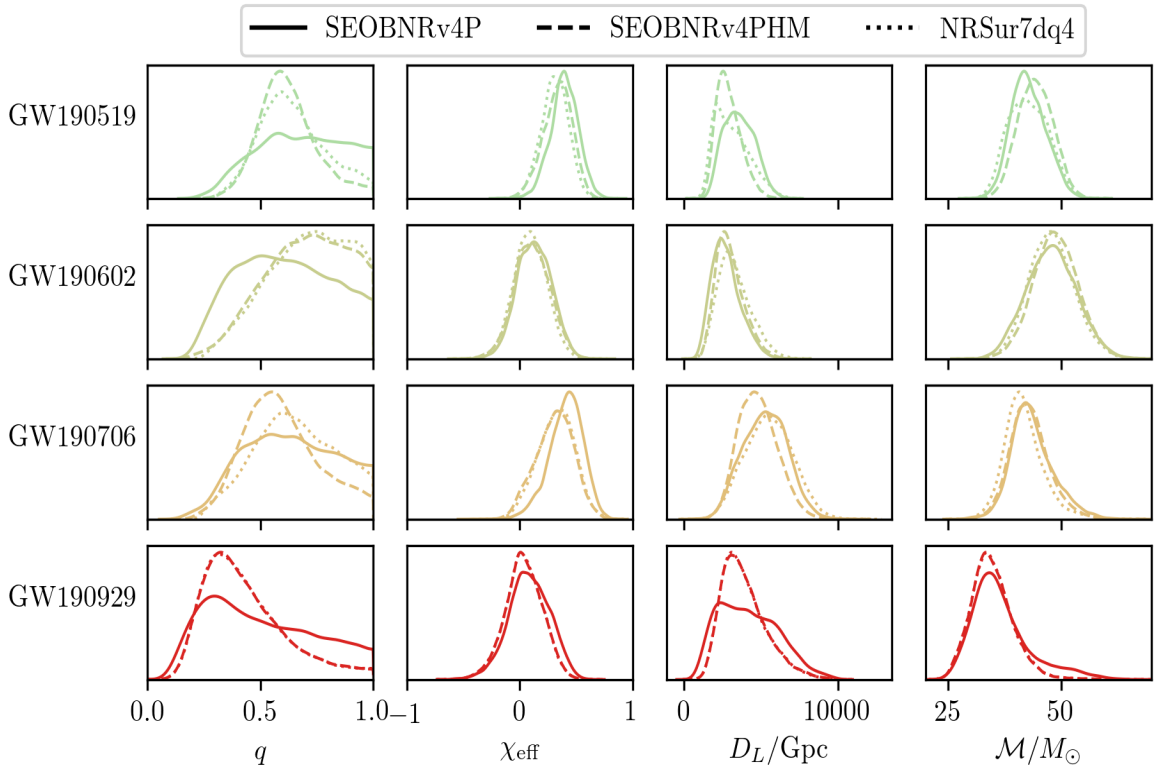

Figure 4.36: **Impact of Higher Order Modes:** This figure shows the panels of the 1D marginal distributions for the mass ratio, $\chi_{\rm eff}$, d_L, and $\mathcal{M}^{source}(M_\odot)$ parameters from the events that had the most dramatic difference when including higher order modes (GW190519, GW190602, GW190706, and GW190929). In all panels, SEOBNRv4PHM, SEOBNRv4P and NRSur7dq4 are the dashed, solid, and dotted lines respectively. To get a more quantifiable answer on the information gained when including higher order modes. see the Bayes Factors in Table 4.2 (see [7] for a similar plot).

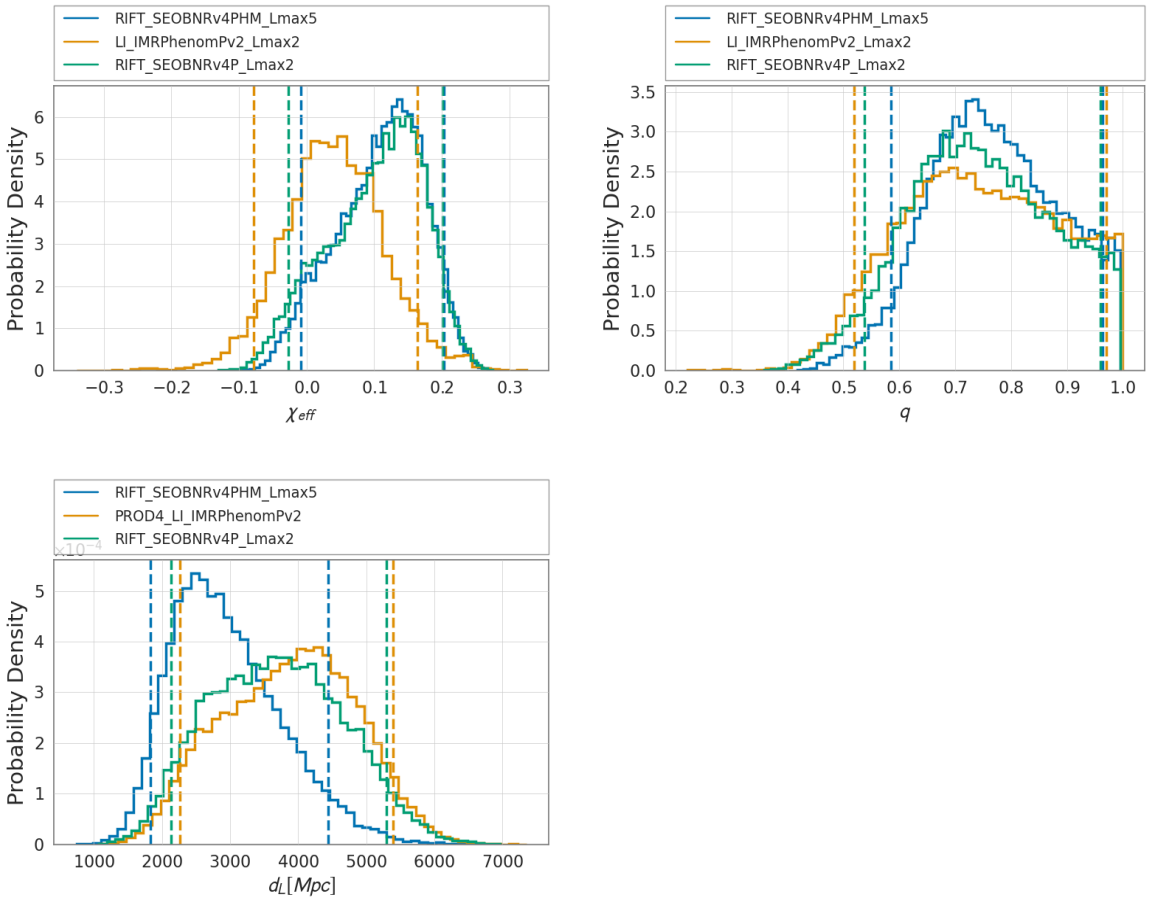

Figure 4.37: **Wavform Systematics on GW190521r:** This figure shows the $\chi_{\rm eff}$, q, and d_L distributions from the different analyses of GW190521r. The results include models SEOBNRv4PHM (blue) with all available modes and SEOBNRv4P (green) done with RIFT and IMRPhenomPv2 (orange) done with LALInference. The JSD between IMRPhenomPv2 and SEOBNRv4P $\chi_{\rm eff}$ distributions is 0.103.

Chapter 5

Conclusions

In this chapter, I summarize the main three subjects of this dissertation: (i) the introduction and validation of a new parameter inference algorithm: RIFT (Section 5.1), (ii) the evaluation of the impact of waveform systematics and HOMs in both synthetic and real GW data (Section 5.2), and (iii) the analyses of real O3a GW data (Section 5.3). We discuss the past and future impact of the field in Section 5.4 and briefly discuss planned future work in Section 5.5.

5.1 RIFT

In Chapter 2, we have introduced and validated RIFT, a strategy to iteratively produce high-precision posterior distributions of binary parameters for a wide variety of candidate compact binary coalescences. We demonstrated through anecdotal and systematic examples how this method could employ costly and even heterogeneous approximations. As concrete illustrations of its utility, we have employed computationally taxing models, which require up to one hour per waveform evaluation, to infer parameters of synthetic binary neutron star systems; We demonstrated this approach reproduces the results of other inference strategies, when employing the same approximations, and provided proof-of-concept demonstrations of new strategies to identify, assess, and remedy systematic errors which our method uniquely enables.

While RIFT does not have the very mature feature set provided by LI, including marginalization over calibration error, this approach is compatible with these extensions, and can

exist simultaneously with other approaches within future parameter inference strategies for gravitational wave sources.

The tools demonstrated in this work are actively being generalized. For example, the gaussian process interpolation strategy can and has been trivially applied to the (fixed) grid provided by numerical relativity simulations, generalizing prior work [9, 129]. Our strategy can also be trivially employed to construct posterior distributions that mix approximations, using different approaches in different parts of parameter space, with the boundaries differentiating between approximations chosen post-facto and with essentially negligible additional computational cost, besides post-processing. Finally, in this work we demonstrated reliable posterior inferences using only $O(10^4)$ physically distinct realizations of GW radiation from a coalescing binary, even for the most complex and high-dimensional binary configurations. As each of the several thousand NR simulations of coalescing binaries can be scaled to arbitrary binary masses, we expect the application RIFT to NR simulations will enable inferences for generic high-mass binary black holes.

While RIFT produced results with dramatically smaller wall-clock time, the investigations performed in this work had overall CPU costs comparable to or even in excess of a corresponding LI analysis with low-cost waveform models. In the future, when overall computational efficiency becomes a more important constraint on our investigations, we will describe optimizations to this strategy which reduce our overall computational burden.

5.2 Significance of waveform and higher order mode systematics

Waveform and HOM systematics significantly impact your ability to infer parameters of GW sources. Using RIFT and other tools, we have been able to make statements about the severity of these systematics when inferring parameters of both synthetic and real sources. In Section 3.2 we first demonstrated via an inner product, which maximizes over t, $\phi_{\rm orb}$, and $\phi_{\rm JL}$, of two commonly used waveform models (IMRPhenomPv2 and SEOBNRv3) generated with identical parameters can often be significantly different from one another. These differences are most evident for systems that are strongly precessing BBHs. When looking at sources

with lower masses and therefore longer signal durations, we found that the waveforms had a greater opportunity to dephase leading to a lower match. To understand how these differences affect the parameter inference, we generated fake sources using both SEOBNRv3 and NR and performed PE with IMRPhenomPv2. While we did find some bias in the inference results, the error we measure in the mismatch did not always manifest itself in the inference. While the inference bias was not always the most extreme, these waveform differences will increasingly affect the inference as the sources become more extreme.

Section 3.3 was part of a larger work that focused on developing new NR simulations from the PE using semi-analytical models. While these simulations agree well with each other, the agreement between the NR simulations and the models were orders of magnitude worse. To illustrate this, we calculated the marginalize log likelihood between the data and each of these new simulations simulation as well as with the models with identical parameters. We found that the marginalize likelihood calculated with NR largely agree between the difference NR codes up to Monte Carlo and fitting error; however, we also found that these likelihood values are substantially lower than the likelihoods computed using the models. Since the likelihood surface is nearly flat near the peak, small differences are amplified in parameter space and insures that any offsets reflect a small change in the likelihood. Alternatively, we chose 5 already existing simulations that produced likelihood values near the peak. We again calculated the likelihood between the data and the models at these new parameters. We again found substantial differences between the NR and model likelihood values; however, here we found the NR likelihoods to be larger. This suggests that the NR and models favor different parts of parameter space and the process used to determine the NR follow up points was not efficient in find the most likely parameters. This work is an example of waveform differences affecting the inferred parameters. These differences might seem minor due to the amplitude of GW170104; however, the effect on the PE will only amplify as louder signals are detected.

In Section 3.4, we used a recently developed NR hybrid surrogate model to thoroughly investigate the impact of HOMs when inferring the parameters of a source. For this study, we use high mass sources with a wide range of mass ratios and spins similar to the sources thought

to be detected by current LVC sensitivities. Previous studies [144, 143, 150, 156, 145, 190, 142, 136, 191, 167, 129, 192] have also explored this question in various approximate contexts; however, here we perform a fully Bayesian three-detector analysis. As expect, we found that HOMs play an important part in accurately inferring the parameters with non-zero spins and $q > 1$ as well as source that had $SNR \geq 30$. Surprisingly, we found that HOMs can noticeably change the recover posteriors even for a low amplitude signal from the simplest source (zero spin; equal mass). Generally HOMs become more important as the source increases in mass ratio and decreases in $\chi_{\rm eff}$. Consistent with [144], we found little to no bias for systems that had large positive spin. Conversely, we found the analyses that omit HOMs were consistently biased away from the true parameters. This can have a significant effect on other analyses such as population inference or tests of general relativity. Any of these analyses that involve high, asymmetric masses will need to appreciate the potentially large waveform systematics that could bias their result. These could lead to inaccurate population inference as well as seemingly failed test of general relativity, see Section 3.4.4.3 and 3.20.

In Section 3.5, we inferred the parameters of the loudest trigger from LIGO's 2015-2017, GW170502, using RIFT with NR surrogate models that include many HOMs. While changes the posterior distributions dramatically, we demonstrated that the impact of HOMs can be noticed even in the analysis for marginal triggers. Using these more sophisticated NR surrogate models, we found the primary and secondary masses to be $\sim 90 M_\odot$ and $\sim 60 M_\odot$ respectively. We found HOMs have a small but noticeable impact on the inference of the inclination, distance, and mass ratio. When including HOMs, the distance distribution shifts toward larger distance while inclinations are more constrained around face on orientations. The mass ratio peaks closer to unity when including the HOMs. As the sensitivity of the detectors increase, these small but noticeable differences caused by the inclusion of HOMs will increase for louder GW170502-like sources.

5.3 Multi-model Analysis of O3a events

In Chapter 4, I analyze 15 binary black holes candidates and 1 binary neutron star candidate. I not only analyzed these events with standard dominant mode-only models, but also models that contain HOMs. I was able to infer all the parameters of the GW source as well as make statements about the importance of waveform and HOMs systematics. The most massive source that was detected was GW190521g with a *detector* frame total mass M^{det} of SEOBNRv4PHM: $276.3^{+41.2}_{-32.6}$; NRSur7dq4: $271.4^{+22.5}_{-24.7}$; NR: $270.5^{+23.1}_{-24.5}$; NR+NRSur7dq4: $270.9^{+23.5}_{-25.9}$. Regardless of which analyses one looks at, these total masses put this event higher than any detected in O1/O2 Observing Runs [15]. The least massive binary is GW190425 (SEOBNRv4T_surrogate low spin: $3.3^{+0.06}_{-0.05}$, high spin: $3.3^{+0.2}_{-0.07}$; TEOBResumS low spin: $3.3^{+0.06}_{-0.05}$, high spin: $3.3^{+0.2}_{-0.08}$). While these masses imply that this event is most likely a binary neutron star, the tidal parameters inferred do not rule out the possibility of this event being a low-mass binary black hole or neutron star-black hole binary.

I found that most events were consistent with a effective spin $\chi_{\rm eff}$ of zero within uncertainties; however, some candidates can be constrained away from zero. The event with the largest positive median effective spin was GW190517 with SEOBNRv4PHM: $0.56^{+0.19}_{-0.18}$; NRSur7dq4 (all modes): $0.50^{+0.18}_{-0.19}$. While no system among the events I analyzed can confidently be constrained to $\chi_{\rm eff} < 0$, the system with the most negative median effective spin was GW190701 (SEOBNRv4PHM: $-0.04^{+0.23}_{-0.30}$; NRSur7dq4 (all modes): $-0.08^{+0.22}_{-0.27}$). To quantify any evidence of precession, I calculated the Bayes Factors between precessing vs non-precessing systems. Most of the Bayes Factors were consistent with 1 (neither favoring precessing nor non-precessing systems); however, there was evidence of mild precession for GW190521g and GW190828ℓ with a corresponding Bayes Factors of 4.38 and 2.28.

The inclusion of HOMs affected all events to varying degrees of severity. The most significant effect can be seen in the distributions of GW190519. When including HOMs, the parameters inferred favored more asymmetric mass ratios and thus a higher source-frame primary mass, edge on orientations, lower spins, and smaller distances. To quantify this the information gained when including HOMs, we calculated a Bayes Factor$\sim$ 15. To a lesser

extent, HOMs excludes more systems with extreme asymmetric and near equal mass ratios for GW190706 and GW190929. Conversely when including the HOMs, GW190602 is more confidently constrained to closer to equal mass systems. While there is visible systematic waveform differences present in the posteriors, they are still secondary to statistical effects.

5.4 Impact of the Field

RIFT was originally developed to resolve the need of fast and accurate PE for the many GW sources expected to appear. Since the very first detection, RIFT has been used to analyze real data. In [9] a version of RIFT was used to analyze GW150914 using waveforms generated NR simulations. In this analysis, the mass ratio was better constrained when including HOMs. This analysis was the first time NR waveforms including HOMs were used to analyze a real GW event. RIFT was also used to validate inferences about GW170608, again via directly comparing GW data to numerical relativity simulations [193]. RIFT also appeared multiple times in the O2 Catalog paper: (i) it was used to reanalyze the binary neutron star merger GW170817 with two more sophisticated models, (ii) it was used a precessing analysis for GW170814, (iii) and it was used to make statements about HM modes, again via directly comparing GW data to numerical relativity simulations[15] (As an aside, RIFT was also used for many unpublished PE investigations with model-based inference, which validated the precessing analyses for most events in O2). RIFT was also used in a study to reanalyze the most massive binary black hole from O2: GW170729 [174]. RIFT contributed to the analysis of GW190521g that will be published soon in [185, 6] (also see Chapter 4). In this paper, RIFT was an essential contributor to the PE reported as well as statements about the evidence of precession and HOMs. Similarly in the very near future, the results from the first half of O3a will be published in [7]. RIFT was used to perform all the analyses with HOMs as well as analyses with more costly models. Because these models contain the most physics, RIFT results will usually be presented as the principal science result for every event in the O3a catalog. In addition, by providing both analyses with and without HOMs, RIFT was essential ingredient into investigations of the effects of HOMs as well as waveform systematics. While

it has been used to varying degrees to analyze real GW data, RIFT continues to be the go-to PE tool when analyzing events with more costly waveforms that include HOMs.

Besides being a major PE code within the LVC, the code can be used in other studies as well. RIFT is currently being used to make statements about binary evolution by evaluating the GW data against models that predict the parameters of the source. The code is also currently being used in multi-messenger inference where the results from the GW analysis is compared to different light curves [194]. The code has also been used by [195] to do a test of general relativity using NR simulations. While other studies do similar test with different tools [171, 196, 197], RIFT allows for a wider range of waveform models to be used. Because waveform systematics could mimic the effect of a breakdown of general relativity, it is important to use a full suite of techniques to analyze GW data when testing the validity of general relativity.

With the ever increasing sensitivities of the LIGO-Virgo observatories, we are fast approaching a period where the waveform systematics will be the dominant source of errors. As waveform models get ever more sophisticated, it is likely that the computational cost of evaluating these models will also increase. RIFT is in an prime position to further our understanding of these waveform systematics and to attempt to resolve them by utilizing all waveforms available regardless of the computational costs.

5.5 Future Work

Continuing the work presented in Section 3.5 and Chapter 4, I plan to continue to analyze real GW signals detected by the LVC. As waveform models become more sophisticated, they will most likely also become more computationally expensive. At the same time, the sensitivity of the LVC detectors constantly being improved and therefore can detect more sources with louder signals. Because of the speed and utility of RIFT, it can play a vital role in analyzing the monumental number of events with the most state-of-the-art sophisticated models available at a given time. Similar to previous work [18] and Section 4.2.6, I plan to continue to analyze signals by directly comparing to NR waveforms. New simulations are always being produced;

further breakthroughs in NR will allow for faster run timescales and the ability to explore the extreme parts of parameter space. This will allow our (fixed) NR grid to expand into more exotic parts of parameter space and in turn allow for analysis of extreme sources with only NR waveforms. As was shown in Section 3.3, models and NR waveforms can favor different parts of parameter space for a given event. Due to this, I hope to reanalyze the known GW detections only by comparing to NR waveforms. I can then attempt to make statements on the impact waveform and HOMs systematics by comparing to other model-based analyses.

As the masses of a source decrease, the length of the waveform required for PE needs to increase. Due to the restrictions of performing a NR simulation, most simulations include 10s of cycles before merger. At a certain point in mass, the length of the signal will be greater than the NR waveform simulation. If we are to have any hope in analyzing low mass events with NR waveforms, we will have to hybridize the merger of our NR waveforms with the inspiral of a model. These hybrids have the benefit of both (i) having an accurate merger-ringdown and (ii) being able to have an arbitrary long inspiral. However, these hybrid waveforms will need to accurately incorporate the precession and HOMs of a given NR waveform. The methodology of this was recently published in [198]. Building off of this work, I plan to incorporate this framework into RIFT to allow for on-the-fly hybridization of a given NR waveform. Besides validating that the PE recovered from using the hybrid waveforms are accurate, I plan to investigate hybrids involving the inspirals from different models. Moving forward, this will guarantee we have the most accurate hybrid possible when analyzing low mass sources.

Building off the work in [18, 3], Chapter 2, and Section 3.3, I plan on developing and validating an algorithm that suggests new NR simulations for a given analysis of an event. Once the grid of likelihood points are evaluated, we construct the continuous likelihood distribution by interpolating the points in the relevant parameters via a Gaussian Process Regression (Remember Section 2.1.3). One additional feature of the Gaussian process, besides outputting an interpolated distribution, is it outputs an error estimate to said interpolation. One might naively think that the NR follow up simulation should be run where the highest error is estimated; however, this can turn out to be in an irrelevant part of parameter space for the

given event being analyzed. Instead I plan to produce NR follow up suggestions based on both places where the error to the fit is high as well as taking into account the parameter space where the likelihood is high for a given event. It will be critical to the study to figure out how much weight should be given to the error and to the likelihood. Once successfully implemented and validated, we can suggest NR follow up points based off of analyses done with NR and models and compare these new simulations (similar to the study in Section 3.3)

Continuing with the work presented in Section 3.2 and 3.4, I plan on expanding studies of waveform systematics and HOMs. Many of the analyses done in 3.2 used state-of-the-art models at the time of the work but have now become eclipsed by newer models that include more information about the signal (See Section 1.2 for examples of these newer models). I plan to use similar tools to study the systematics between these newer, more complete models. With the inclusion of RIFT in this future study, I will be able to obtained a more complete analysis that is not hindered by unrealistic run timescales. I also plan to continue our synthetic studies of the effects of HOMs. Specifically, I plan on expanding our results presented in Section 3.4 to include precession. The release of new, precessing NR surrogate models will allow us to investigate the impact of HOMs when measuring the individual spins as well as precession.

As the next era of GW astronomy approaches, new observatories such as the Laser Interferometer Space Antenna (LISA) [199] plan to come online. This space-based GW observatory will allow for detections of GWs with different frequencies and therefore sources. While the waveforms for PE needed for LISA are similar to the ones used by the LVC, the accuracy needed for the waveforms are significantly larger; therefore, studies like the ones presented in this dissertation will be essential to understanding the systematics introduced while analyzing LISA sources.

Bibliography

[1] Jacob Lange, Richard O'Shaughnessy, and Monica Rizzo. Rapid and accurate parameter inference for coalescing compact binaries. *Submitted to PRD*, 2019. (document), 1.1.1, 2, 3.4.2

[2] A. R. Williamson, J. Lange, R. O'Shaughnessy, J. A. Clark, Prayush Kumar, J. Calderón Bustillo, and J. Veitch. Systematic challenges for future gravitational wave measurements of precessing binary black holes. *Phys. Rev.*, D96(12):124041, 2017. (document), 2.3.2, 2.4.3, 3.1, 3.4.1

[3] J. Healy, J. Lange, R. O'Shaughnessy, C. O. Lousto, M. Campanelli, A. R. Williamson, Y. Zlochower, J. Calderón Bustillo, J. A. Clark, C. Evans, D. Ferguson, S. Ghonge, K. Jani, B. Khamesra, P. Laguna, D. M. Shoemaker, M. Boyle, A. García, D. A. Hemberger, L. E. Kidder, P. Kumar, G. Lovelace, H. P. Pfeiffer, M. A. Scheel, and S. A. Teukolsky. Targeted numerical simulations of binary black holes for GW170104. *Phys. Rev. D*, 97:064027, Mar 2018. (document), 3.1, 3.2, 3.3, 3.3.1, 3.3.1.1, 3.11, 3.3.2, 3.3.2, 5.5

[4] Feroz H. Shaik, Jacob Lange, Scott E. Field, Richard O'Shaughnessy, Vijay Varma, Lawrence E. Kidder, Harald P. Pfeiffer, and Daniel Wysocki. Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems. *Phys. Rev. D*, 101(12):124054, June 2020. (document), 3.1, 3.6

[5] Richard Udall, Karan Jani, Jacob Lange, Richard O'Shaughnessy, James Clark, Laura Cadonati, Deirdre Shoemaker, and Kelly Holley-Bockelmann. Inferring parameters of

the loudest intermediate mass black hole trigger in LIGO's O1/O2 data. *arXiv e-prints*, page arXiv:1912.10533, December 2019. (document), 3.1, 3.5

[6] LIGO Scientific Collaboration and Virgo Collaboration. Properties and astrophysical implications of the 150 $M_{\odot}$ binary black hole merger GW190521.1. Technical Report P2000021, LSC, 2020. (document), 1.2, 4.1, 4.2, 4.2.1, 4.2.6, 5.4

[7] LIGO Scientific Collaboration and Virgo Collaboration. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. Technical Report P2000061, LSC, 2020. (document), 1.2, 4.1, 4.2, 4.36, 5.4

[8] Daniel Foreman-Mackey. corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 24, 2016. (document), 2.1

[9] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration). Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence. *Phys. Rev. D*, 94:064035, Sep 2016. (document), 2.1.2, 2.1.5, 2.3, 2.3.1, 2.5, 2.5, 3.2, 3.2.1.2, 3.1, 3.5, 3.6, 3.8, 3.2.3.1, 3.3, 3.3.2, 5.1, 5.4

[10] B. P. Abbott et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.*, X6(4):041015, 2016. [Erratum: Phys. Rev. X8, no.3, 039903(2018)]. (document), 2.5, 3.2.1.1, 3.5, 3.2.2, 3.2.3.2

[11] S. Babak, A. Taracchini, and A. Buonanno. Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity. *Phys. Rev. D*, 95(2):024010, January 2017. (document), 1.2, 3.7, 3.3.1.1

[12] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev. D*, 93(4):044007, February 2016. (document), 1.2, 3.7

[13] Geoffrey Lovelace, Carlos O. Lousto, James Healy, Mark A. Scheel, Alyssa Garcia, Richard O'Shaughnessy, Michael Boyle, Manuela Campanelli, Daniel A. Hemberger,

Lawrence E. Kidder, Harald P. Pfeiffer, Béla Szilágyi, Saul A. Teukolsky, and Yosef Zlochower. Modeling the source of GW150914 with targeted numerical-relativity simulations. *Classical and Quantum Gravity*, 33(24):244002, December 2016. (document), 3.3.1.1, 3.10

[14] Yosef Zlochower and Carlos O. Lousto. Modeling the remnant mass, spin, and recoil from unequal-mass, precessing black-hole binaries: The Intermediate Mass Ratio Regime. *Phys. Rev.*, D92:024022, 2015. (document), 3.5, 3.3.2, 3.12

[15] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev.*, X9(3):031040, 2019. (document), 2.1.7, 3.4.3, 3.4.3.4, 3.4.4.2, 3.4.4.4, 3.24, 3.5.1, 4.2.1, 4.2.3, 5.3, 5.4

[16] Barak Zackay, Liang Dai, Tejaswi Venumadhav, Javier Roulet, and Matias Zaldarriaga. Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers. *arXiv e-prints*, page arXiv:1910.09528, Oct 2019. (document), 3.5, 3.24

[17] Alexander H. Nitz, Thomas Dent, Gareth S. Davies, Sumit Kumar, Collin D. Capano, Ian Harry, Simone Mozzon, Laura Nuttall, Andrew Lundgren, and Márton Tápai. 2-OGC: Open Gravitational-wave Catalog of Binary Mergers from Analysis of Public Advanced LIGO and Virgo Data. *Astrophysical Journal*, 891(2):123, March 2020. (document), 3.5, 3.24

[18] J. Lange, R. O'Shaughnessy, M. Boyle, J. Calderón Bustillo, M. Campanelli, T. Chu, J. A. Clark, N. Demos, H. Fong, J. Healy, D. A. Hemberger, I. Hinder, K. Jani, B. Khamesra, L. E. Kidder, P. Kumar, P. Laguna, C. O. Lousto, G. Lovelace, S. Ossokine, H. Pfeiffer, M. A. Scheel, D. M. Shoemaker, B. Szilagyi, S. Teukolsky, and Y. Zlochower. Parameter estimation method that directly compares gravitational wave observations to numerical relativity. *Phys. Rev. D*, 96(10):104041, November 2017. (document), 1, 1.1.1, 1.2, 1.3.2, 1.3.2, 2.1.2, 3.3, 3.3.2, 4.1, 5.5

[19] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016. 1, 2.5

[20] B. Abbott et al. (The LIGO Scientific Collaboration). Advanced LIGO. *Class. Quant. Grav.* , 32(7):074001, April 2015. 1

[21] T. Accadia et al. Virgo: a laser interferometer to detect gravitational waves. *JINST*, 7:P03012, 2012. 1

[22] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy. Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences. *Phys. Rev.*, D92(2):023002, 2015. 1, 1.1.1, 1.1.1, 1.3.1, 1.3.1.2, 2.1.2, 2.1.2, 2.1.4, 2.1.5, 2.6.1, 3.3.2, 3.4.2

[23] J. Veitch, V. Raymond, B. Farr, W. M. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, W. D. Pozzo, F. Feroz, J. Gair, C. Haster, V. Kalogera, T. Littenberg, I. Mandel, R. O'Shaughnessy, M. Pitkin, C. Rodriguez, C. Röver, T. Sidery, R. Smith, M. V. D. Sluys, A. Vecchio, W. Vousden, and L. Wade. Robust parameter estimation for compact binaries with ground-based gravitational-wave observations using LALInference. *Phys. Rev. D*, 91:042003, Feb 2015. 1.1.1, 2.1.7, 2.6.2, 3.2.1.2, 3.2.2, 3.3.1

[24] T. Damour. Coalescence of two spinning black holes: An effective one-body approach. *Phys. Rev. D*, 64(12):124013, December 2001. 1.1.2

[25] É. Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. *Phys. Rev. D*, 78(4):044021, August 2008. 1.1.2, 3.2.3.2

[26] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Bruegmann, et al. Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins. *Phys. Rev. Lett.*, 106:241101, 2011. 1.1.2, 3.2.3.2, 3.5.1

[27] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel. Prototype effective-one-body model for nonprecessing

spinning inspiral-merger-ringdown waveforms. *Phys. Rev. D*, 86(2):024011, July 2012. 1.2

[28] C. M. Will and A. G. Wiseman. Gravitational radiation from compact binary systems: Gravitational waveforms and energy loss to second post-Newtonian order. *Phys. Rev. D*, 54:4813–4848, October 1996. 1.2

[29] L. E. Kidder. Coalescing binary systems of compact objects to $(\text{post})^{5/2}$-Newtonian order. V. Spin effects. *Phys. Rev. D*, 52:821–847, July 1995. 1.2

[30] Theocharis A. Apostolatos, Curt Cutler, Gerald J. Sussman, and Kip S. Thorne. Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries. *Phys. Rev. D*, 49:6274, 1994. 1.2

[31] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri. A physical template family for gravitational waves from precessing binaries of spinning compact objects: Application to single-spin binaries. *Phys. Rev. D*, 69(10):104017–+, May 2004. 1.2

[32] P. Schmidt, M. Hannam, S. Husa, and P. Ajith. Tracking the precession of compact binaries from their gravitational-wave signal. *Phys. Rev. D*, 84(2):024046, July 2011. 1.2

[33] L. Pekowsky, R. O'Shaughnessy, J. Healy, and D. Shoemaker. Comparing gravitational waves from nonprecessing and precessing black hole binaries in the corotating frame. *Phys. Rev. D*, 88:024040, Jul 2013. 1.2

[34] M. Boyle. Angular velocity of gravitational radiation from precessing binaries and the corotating frame. *Phys. Rev. D*, 87(10):104006, May 2013. 1.2

[35] R. O'Shaughnessy, L. London, J. Healy, and D. Shoemaker. Precession during merger: Strong polarization changes are observationally accessible features of strong-field gravity during binary black hole merger. *Phys. Rev. D*, 87(4):044038, February 2013. 1

[36] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi. Numerical relativity waveform surrogate model

for generically precessing binary black hole mergers. *Phys. Rev. D*, 96(2):024058, July 2017. 1, 1.2, 2.2, 3.2, 3.2.3.2

[37] S. Ossokine, M. Boyle, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi. Comparing post-Newtonian and numerical relativity precession dynamics. *Phys. Rev. D*, 92(10):104028, November 2015. 1.2

[38] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. *Phys. Rev. D*, 95(4):044028, February 2017. 1.2, 2.3.1

[39] Andrea Taracchini et al. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys. Rev.*, D89(6):061502, 2014. 1.2, 2.2, 3.2.1.1, 3.4.4.5

[40] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel. Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism. *Phys. Rev. D*, 84(12):124052, December 2011. 1.2

[41] Enrico Barausse, Alessandra Buonanno, Scott A. Hughes, Gaurav Khanna, Stephen O'Sullivan, and Yi Pan. Modeling multipolar gravitational-wave emission from small mass-ratio mergers. *Phys. Rev. D*, 85(2):024046, January 2012. 1.2

[42] Roberto Cotesta, Alessandra Buonanno, Alejandro Bohé, Andrea Taracchini, Ian Hinder, and Serguei Ossokine. Enriching the symphony of gravitational waves from binary black holes by tuning higher harmonics. *Phys. Rev. D*, 98(8):084028, October 2018. 1.2, 3.4.1

[43] Serguei Ossokine, Alessandra Buonanno, Sylvain Marsat, Roberto Cotesta, Stanislav Babak, Tim Dietrich, Roland Haas, Ian Hinder, Harald P. Pfeiffer, Michael Pürrer, Charles J. Woodford, Michael Boyle, Lawrence E. Kidder, Mark A. Scheel, and Béla

Szilágyi. Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construction and Validation. *arXiv e-prints*, page arXiv:2004.09442, April 2020. 1.2

[44] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. T. Whelan, B. Brügmann, P. Diener, J. Gonzalez, M. Hannam, S. Husa, M. Koppitz, D. Pollney, L. Rezzolla, L. Santamaría, A. M. Sintes, U. Sperhake, and J. Thornburg. A phenomenological template family for black-hole coalescence waveforms. *Classical and Quantum Gravity*, 24:689–+, October 2007. 1.2

[45] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan. Matching post-Newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. *Phys. Rev. D*, 82(6):064016, September 2010. 1.2

[46] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. *Phys. Rev. Lett.*, 113(15):151101, 2014. 1.2, 2.2, 3.2.1.1, 4.2.2

[47] Lionel London, Sebastian Khan, Edward Fauchon-Jones, Cecilio García, Mark Hannam, Sascha Husa, Xisco Jiménez-Forteza, Chinmay Kalaghatgi, Frank Ohme, and Francesco Pannarale. First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries. *Phys. Rev. Lett*, 120(16):161102, April 2018. 1.2, 3.4.1

[48] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. *Phys. Rev. D*, 100(2):024059, July 2019. 1.2

[49] Sebastian Khan, Frank Ohme, Katerina Chatziioannou, and Mark Hannam. Including higher order multipoles in gravitational-wave models for precessing binary black holes. *Phys. Rev. D*, 101(2):024056, January 2020. 1.2

[50] Geraint Pratten, Sascha Husa, Cecilio Garcia-Quiros, Marta Colleoni, Antoni Ramos-Buades, Hector Estelles, and Rafel Jaume. Setting the cornerstone for the IMRPhenomX family of models for gravitational waves from compact binaries: The dominant harmonic for non-precessing quasi-circular black holes. *arXiv e-prints*, page arXiv:2001.11412, January 2020. 1.2

[51] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa. Let's twist again: computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes. *arXiv e-prints*, page arXiv:2004.06503, April 2020. 1.2

[52] Cecilio García-Quirós, Marta Colleoni, Sascha Husa, Héctor Estellés, Geraint Pratten, Antoni Ramos-Buades, Maite Mateu-Lucena, and Rafel Jaume. IMRPhenomXHM: A multi-mode frequency-domain model for the gravitational wave signal from non-precessing black-hole binaries. *arXiv e-prints*, page arXiv:2001.10914, January 2020. 1.2

[53] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio. Fast Prediction and Evaluation of Gravitational Waveforms Using Surrogate Models. *Physical Review X*, 4(3):031006, July 2014. 1.2, 3.4.1

[54] M. Pürrer. Frequency-domain reduced order models for gravitational waves from aligned-spin compact binaries. *Classical and Quantum Gravity*, 31(19):195010, October 2014. 1.2, 3.4.4.5

[55] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger. Fast and Accurate Prediction of Numerical Relativity Waveforms from Binary Black Hole Coalescences Using Surrogate Models. *Physical Review Letters*, 115(12):121102, September 2015. 1.2, 2.2

[56] B. D. Lackey, S. Bernuzzi, C. R. Galley, J. Meidam, and C. Van Den Broeck. Effective-one-body waveforms for binary neutron stars using surrogate models. *Phys. Rev. D*, 95(10):104036, May 2017. 1.2

[57] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Lawrence E. Kidder, and Harald P. Pfeiffer. Surrogate model of hybridized numerical relativity binary black hole waveforms. *Phys. Rev.*, D99(6):064045, 2019. 1.2, 3.4.1, 3.5

[58] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith. A Surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary black hole mergers. *Phys. Rev. D*, 95(10):104023, May 2017. 1.2, 2.2, 3.2, 3.2.3.2

[59] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt. Fast and accurate inference on gravitational waves from precessing compact binaries. *Phys. Rev. D*, 94(4):044031, August 2016. 1.2

[60] É. É. Flanagan and T. Hinderer. Constraining neutron-star tidal Love numbers with gravitational-wave detectors. *Phys. Rev. D*, 77(2):021502, January 2008. 1.2

[61] J. Vines, É. É. Flanagan, and T. Hinderer. Post-1-Newtonian tidal effects in the gravitational waveform from binary inspirals. *Phys. Rev. D*, 83(8):084051, April 2011. 1.2

[62] T. Dietrich, S. Bernuzzi, and W. Tichy. Closed-form tidal approximants for binary neutron star gravitational waveforms constructed from high-resolution numerical relativity simulations. *Phys. Rev. D*, 96(12):121501, December 2017. 1.2

[63] Tim Dietrich, Sebastian Khan, Reetika Dudi, Shasvath J. Kapadia, Prayush Kumar, Alessandro Nagar, Frank Ohme, Francesco Pannarale, Anuradha Samajdar, Sebastiano Bernuzzi, Gregorio Carullo, Walter Del Pozzo, Maria Haney, Charalampos Markakis, Michael Pürrer, Gunnar Riemenschneider, Yoshinta Eka Setyawati, Ka Wa Tsang, and Chris Van Den Broeck. Matter imprints in waveform models for neutron star binaries: Tidal and self-spin effects. *Phys. Rev. D*, 99(2):024029, January 2019. 1.2, 2.3.3

[64] Alessandro Nagar, Sebastiano Bernuzzi, Walter Del Pozzo, Gunnar Riemenschneider, Sarp Akcay, Gregorio Carullo, Philipp Fleig, Stanislav Babak, Ka Wa Tsang, Marta Colleoni, Francesco Messina, Geraint Pratten, David Radice, Piero Rettegno, Michalis Agathos, Edward Fauchon-Jones, Mark Hannam, Sascha Husa, Tim Dietrich, Pablo Cerdá-Duran, José A. Font, Francesco Pannarale, Patricia Schmidt, and Thibault Damour. Time-domain effective-one-body gravitational waveforms for coalescing compact binaries with nonprecessing spins, tides, and self-spin effects. *Phys. Rev. D*, 98(10):104052, November 2018. 1.2

[65] S. Bernuzzi, A. Nagar, T. Dietrich, and T. Damour. Modeling the Dynamics of Tidally Interacting Binary Neutron Stars up to the Merger. *Physical Review Letters*, 114(16):161103, April 2015. 1.2

[66] T. Hinderer, A. Taracchini, F. Foucart, A. Buonanno, J. Steinhoff, M. Duez, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, B. Szilagyi, K. Hotokezaka, K. Kyutoku, M. Shibata, and C. W. Carpenter. Effects of Neutron-Star Dynamic Tides on Gravitational Waveforms within the Effective-One-Body Approach. *Physical Review Letters*, 116(18):181101, May 2016. 1.2

[67] J. Steinhoff, T. Hinderer, A. Buonanno, and A. Taracchini. Dynamical tides in general relativity: Effective action and effective-one-body Hamiltonian. *Phys. Rev. D*, 94(10):104028, November 2016. 1.2

[68] S. Marsat and J. Vines. SEOBNRv2T/v4T update: inclusion of quadrupole-monopole terms. *LIGO Technical Note T1800028, available at https://dcc.ligo.org/LIGO-T1800028*, January 2018. 1.2

[69] Benjamin D. Lackey, Michael Pürrer, Andrea Taracchini, and Sylvain Marsat. Surrogate model for an aligned-spin effective-one-body waveform model of binary neutron star inspirals using Gaussian process regression. *Phys. Rev. D*, 100(2):024002, July 2019. 1.2

[70] E. Poisson. Gravitational waves from inspiraling compact binaries: The quadrupole-moment term. *Phys. Rev. D*, 57:5287–5290, April 1998. 1.2

[71] Frans Pretorius. Evolution of binary black hole spacetimes. *Phys. Rev. Lett.*, 95:121101, 2005. 1.2

[72] Manuela Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower. Accurate evolutions of orbiting black-hole binaries without excision. *Phys. Rev. Lett.*, 96:111101, 2006. 1.2

[73] John G. Baker, Joan Centrella, Dae-Il Choi, Michael Koppitz, and James van Meter. Gravitational wave extraction from an inspiraling configuration of merging black holes. *Phys. Rev. Lett.*, 96:111102, 2006. 1.2

[74] Abdul H. Mroue et al. Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy. *Phys. Rev. Lett.*, 111(24):241104, 2013. 1.2, 2.2, 3.2.2

[75] Karan Jani, James Healy, James A. Clark, Lionel London, Pablo Laguna, and Deirdre Shoemaker. Georgia Tech Catalog of Gravitational Waveforms. *Class. Quant. Grav.*, 33(20):204001, 2016. 1.2

[76] J. Healy, C. O. Lousto, Y. Zlochower, and M. Campanelli. The RIT binary black hole simulations catalog. *Classical and Quantum Gravity*, 34(22):224001, November 2017. 1.2, 2.2

[77] James Healy, Carlos O. Lousto, Jacob Lange, Richard O'Shaughnessy, Yosef Zlochower, and Manuela Campanelli. Second RIT binary black hole simulations catalog and its application to gravitational waves parameter estimation. *Phys. Rev. D*, 100(2):024021, July 2019. 1.2

[78] Michael Boyle et al. The SXS Collaboration catalog of binary black hole simulations. *Class. Quant. Grav.*, 36(19):195006, 2019. 1.2, 3.4.1

[79] James Healy and Carlos O. Lousto. The Third RIT binary black hole simulations catalog. 7 2020. 1.2

[80] Y. Zlochower, J.G. Baker, Manuela Campanelli, and C.O. Lousto. Accurate black hole evolutions by fourth-order numerical relativity. *Phys. Rev. D*, 72:024021, 2005. 1.2

[81] F. Löffler, J. Faber, E. Bentivegna, T. Bode, P. Diener, R. Haas, I. Hinder, B. C. Mundim, C. D. Ott, E. Schnetter, G. Allen, M. Campanelli, and P. Laguna. The Einstein Toolkit: a community computational infrastructure for relativistic astrophysics. *Class. Quantum Grav.*, 29(11):115001, June 2012. 1.2

[82] Einstein Toolkit home page: `http://einsteintoolkit.org`. 1.2

[83] Cactus Computational Toolkit home page: `http://cactuscode.org`. 1.2

[84] Erik Schnetter, Scott H. Hawley, and Ian Hawke. Evolutions in 3D numerical relativity using fixed mesh refinement. *Class. Quant. Grav.*, 21:1465–1488, 21 March 2004. 1.2

[85] Frank Herrmann, Ian Hinder, Deirdre M. Shoemaker, Pablo Laguna, and Richard A. Matzner. Binary Black Holes: Spin Dynamics and Gravitational Recoil. *Phys. Rev.*, D76:084032, 2007. 1.2

[86] Frank Herrmann, Ian Hinder, Deirdre Shoemaker, Pablo Laguna, and Richard A. Matzner. Gravitational recoil from spinning binary black hole mergers. *Astrophys. J.*, 661:430–436, 2007. 1.2

[87] Ian Hinder, Birjoo Vaishnav, Frank Herrmann, Deirdre Shoemaker, and Pablo Laguna. Universality and Final Spin in Eccentric Binary Black Hole Inspirals. *Phys. Rev. D*, 77:081502, 2008. 1.2

[88] James Healy, Frank Herrmann, Ian Hinder, Deirdre M. Shoemaker, Pablo Laguna, et al. Superkicks in Hyperbolic Encounters of Binary Black Holes. *Phys.Rev.Lett.*, 102:041101, 2009. 1.2

[89] Ian Hinder, Frank Herrmann, Pablo Laguna, and Deirdre Shoemaker. Comparisons of eccentric binary black hole simulations with post-Newtonian models. *Phys. Rev. D*, 82(2):024033, July 2010. 1.2

[90] James Healy, Janna Levin, and Deirdre Shoemaker. Zoom-Whirl Orbits in Black Hole Binaries. *Phys. Rev. Lett.*, 103:131101, 2009. 1.2

[91] James Healy, Pablo Laguna, Richard A. Matzner, and Deirdre M. Shoemaker. Final Mass and Spin of Merged Black Holes and the Golden Black Hole. *Phys. Rev.*, D81:081501, 2010. 1.2

[92] Tanja Bode, Roland Haas, Tamara Bogdanovic, Pablo Laguna, and Deirdre Shoemaker. Relativistic Mergers of Supermassive Black Holes and their Electromagnetic Signatures. *Astrophys. J.*, 715, 2010. 1.2

[93] `http://www.black-holes.org/SpEC.html`. 1.2, 2.2

[94] Jacob A. Lange. Directly comparing synthetic and real binary black hole coalescence sources with numerical solutions of einstein's equation. Master's thesis, 2016. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2019-10-18. 1.3.2

[95] R. O'Shaughnessy, J. Blackman, and S. E. Field. An architecture for efficient gravitational wave parameter estimation with multimodal linear surrogate models. *Classical and Quantum Gravity*, 34(14):144002, July 2017. 2.1.2, 2.1.3, 2.6.1, 1, 3.2.3.1, 3.2.3.2

[96] C.E. Rasmussen and C.I.K. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006. 2.1.3

[97] R. van Haasteren and M. Vallisneri. New advances in the Gaussian-process approach to pulsar-timing data analysis. *Phys. Rev. D*, 90(10):104012, November 2014. 2.1.3

[98] C. J. Moore and J. R. Gair. Novel Method for Incorporating Model Uncertainties into Gravitational Wave Parameter Estimates. *Physical Review Letters*, 113(25):251101, December 2014. 2.1.3

[99] Z. Doctor, B. Farr, D. E. Holz, and M. Pürrer. Statistical gravitational waveform models: What to simulate next? *Phys. Rev. D*, 96(12):123011, December 2017. 2.1.3

[100] C. J. Moore, C. P. L. Berry, A. J. K. Chua, and J. R. Gair. Improving gravitational-wave parameter estimation using Gaussian process regression. *Phys. Rev. D*, 93(6):064001, March 2016. 2.1.3

[101] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 2.1.3

[102] D. Wysocki, R. O'Shaughnessy, J. Lange, and Y.-L. L. Fang. Accelerating parameter inference with graphics processing units. *Phys. Rev. D*, 99(8):084026, April 2019. 2.1.7, 3.4.2

[103] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al. Properties of the Binary Neutron Star Merger GW170817. *Physical Review X*, 9(1):011001, January 2019. 2.1.7, 4.2.1

[104] C. Devine, Z. B. Etienne, and S. T. McWilliams. Optimizing spinning time-domain gravitational waveforms for advanced LIGO data analysis. *Classical and Quantum Gravity*, 33(12):125025, June 2016. 2.1.7

[105] A. Nagar and P. Rettegno. Rush the inspiral: efficient Effective One Body time-domain gravitational waveforms. *ArXiv e-prints*, May 2018. 2.1.7

[106] Yi Pan, Alessandra Buonanno, Andrea Taracchini, Lawrence E. Kidder, Abdul H. Mroué, Harald P. Pfeiffer, Mark A. Scheel, and Béla Szilágyi. Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism. *Phys. Rev.*, D89(8):084006, 2014. 2.2, 3.2.1.1

[107] LIGO Scientific Collaboration. LIGO Algorithm Library - LALSuite. free software (GPL), 2018. 2.2

[108] SXS Collaboration. Sxs surrogates. *Available from https://data.black-holes.org/surrogates/index.html.* 2.2

[109] S. Nagar, S. Bernuzzi, and W. Del Pozzo. Effective-one-body model with spin and tidal interactions. *Available at https://bitbucket.org/eob_ihes/teobresums.* 2.2

[110] Prayush Kumar, Tony Chu, Heather Fong, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Mark A. Scheel, and Bela Szilagyi. Accuracy of binary black hole waveform models for aligned-spin binaries. *Phys. Rev. D*, 93(10):104050, May 2016. 2.3.1, 3.2

[111] Daniel Wysocki, Jacob Lange, and Richard O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. *Phys. Rev. D*, 100(4):043012, August 2019. 2.4.3, 3.4.4.3

[112] B. P. Abbott, et. al., (The LIGO Scientific Collaboration, and The Virgo Collaboration). Model comparison from LIGO-Virgo data on GW170817's binary components and consequences for the merger remnant. *Classical and Quantum Gravity*, 37(4):045006, February 2020. 2.4.4

[113] L. Lindblom. Spectral representations of neutron-star equations of state. *Phys. Rev. D*, 82(10):103011, November 2010. 2.4.4

[114] L. Lindblom and N. M. Indik. Spectral approach to the relativistic inverse stellar structure problem II. *Phys. Rev. D*, 89(6):064003, March 2014. 2.4.4

[115] A. Akmal, V. R. Pandharipande, and D. G. Ravenhall. Equation of state of nucleon matter and neutron star structure. *Phys. Rev. c*, 58:1804–1828, September 1998. 2.4.4

[116] P. B. Demorest, T. Pennucci, S. M. Ransom, M. S. E. Roberts, and J. W. T. Hessels. A two-solar-mass neutron star measured using Shapiro delay. *Nature*, 467:1081–1083, October 2010. 2.4.4

[117] J. Antoniadis, P. C. C. Freire, N. Wex, T. M. Tauris, R. S. Lynch, M. H. van Kerkwijk, M. Kramer, C. Bassa, V. S. Dhillon, T. Driebe, J. W. T. Hessels, V. M. Kaspi, V. I. Kondratiev, N. Langer, T. R. Marsh, M. A. McLaughlin, T. T. Pennucci, S. M. Ransom, I. H. Stairs, J. van Leeuwen, J. P. W. Verbiest, and D. G. Whelan. A massive pulsar in a compact relativistic binary. *Science*, 340:448, April 2013. 2.4.4

[118] Matthew F. Carney, Leslie E. Wade, and Burke S. Irwin. Comparing two models for measuring the neutron star equation of state from gravitational-wave signals. *Phys. Rev. D*, 98(6):063004, September 2018. 2.4.4

[119] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens. The LIGO Open Science Center. In *Journal of Physics Conference Series*, volume 610 of *Journal of Physics Conference Series*, page 012021, May 2015. 2.5

[120] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration). Properties of the binary black hole merger GW150914. *Phys. Rev. Lett*, 116:241102, Jun 2016. 2.5, 3.2, 3.2.1.1, 3.2.1.2, 3.2.2

[121] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration). Improved analysis of GW150914 using a fully spin-precessing waveform model. *Phys. Rev. X*, 6:041014, October 2016. 2.5, 3.2.1.1

[122] B. Farr, E. Ochsner, W. M. Farr, and R. O'Shaughnessy. A more effective coordinate system for parameter estimation of precessing compact binaries from gravitational waves. *Phys. Rev. D*, 90(2):024018, July 2014. 2.6.2

[123] J Veitch. Add option for volumetric spin priors. *Private communication*, January 2017. 2.6.2

[124] I. Hinder, A. Buonanno, M. Boyle, Z. B. Etienne, J. Healy, N. K. Johnson-McDaniel, A. Nagar, H. Nakano, Y. Pan, H. P. Pfeiffer, M. Pürrer, C. Reisswig, M. A. Scheel, E. Schnetter, U. Sperhake, B. Szilágyi, W. Tichy, B. Wardell, A. Zenginoğlu, D. Alic, S. Bernuzzi, T. Bode, B. Brügmann, L. T. Buchman, M. Campanelli, T. Chu, T. Damour,

J. D. Grigsby, M. Hannam, R. Haas, D. A. Hemberger, S. Husa, L. E. Kidder, P. Laguna, L. London, G. Lovelace, C. O. Lousto, P. Marronetti, R. A. Matzner, P. Mösta, A. Mroué, D. Müller, B. C. Mundim, A. Nerozzi, V. Paschalidis, D. Pollney, G. Reifenberger, L. Rezzolla, S. L. Shapiro, D. Shoemaker, A. Taracchini, N. W. Taylor, S. A. Teukolsky, M. Thierfelder, H. Witek, and Y. Zlochower. Error-analysis and comparison to analytical models of numerical waveforms produced by the NRAR Collaboration. *Classical and Quantum Gravity*, 31(2):025012, January 2013. 3.2

[125] P. Kumar, K. Barkett, S. Bhagwat, N. Afshari, D. A. Brown, G. Lovelace, M. A. Scheel, and B. Szilágyi. Accuracy and precision of gravitational-wave models of inspiraling neutron star-black hole binaries with spin: Comparison with matter-free numerical relativity in the low-frequency regime. *Phys. Rev. D*, 92(10):102001, November 2015. 3.2

[126] K. Chatziioannou, N. Cornish, A. Klein, and N. Yunes. Spin-precession: Breaking the Black Hole-Neutron Star Degeneracy. *Astrophysical Journal*, 798:L17, January 2015. 3.2

[127] B. Farr, C. P. L. Berry, W. M. Farr, C.-J. Haster, H. Middleton, K. Cannon, P. B. Graff, C. Hanna, I. Mandel, C. Pankow, L. R. Price, T. Sidery, L. P. Singer, A. L. Urban, A. Vecchio, J. Veitch, and S. Vitale. Parameter Estimation on Gravitational Waves from Neutron-star Binaries with Spinning Components. *Astrophysical Journal*, 825:116, July 2016. 3.2

[128] Salvatore Vitale, Davide Gerosa, Carl-Johan Haster, Katerina Chatziioannou, and Aaron Zimmerman. Impact of Bayesian Priors on the Characterization of Binary Black Hole Coalescences. *Phys. Rev. Lett*, 119(25):251103, December 2017. 3.2

[129] B. P. Abbott et al. Effects of waveform model systematics on the interpretation of GW150914. *Class. Quant. Grav.*, 34(10):104002, 2017. 3.2, 3.4.1, 3.4.2, 3.4.3.1, 3.4.3.2, 5.1, 5.2

[130] B. P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017. [Erratum: Phys. Rev. Lett. 121, no.12, 129901(2018)]. 3.2, 3.2.1.1, 3.2.3.2, 3.3.1

[131] D. W. Hogg, A. D. Myers, and J. Bovy. Inferring the Eccentricity Distribution. *Astrophysical Journal*, 725:2166–2175, December 2010. 3.2.1.2

[132] W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio. Distinguishing spin-aligned and isotropic black hole populations with gravitational waves. *Nature*, 548:426–429, August 2017. 3.2.1.2

[133] T. Damour, B. R. Iyer, and B. S. Sathyaprakash. Improved filters for gravitational waves from inspiraling compact binaries. *Phys. Rev. D*, 57:885–907, January 1998. 3.2.1.3

[134] Rainer Storn and Kenneth Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, Dec 1997. 3.2.1.3

[135] R. O'Shaughnessy, B. Farr, E. Ochsner, H.-S. Cho, C. Kim, and C.-H. Lee. Parameter estimation of gravitational waves from nonprecessing black hole-neutron star inspirals with higher harmonics: Comparing Markov-chain Monte Carlo posteriors to an effective Fisher matrix. *Phys. Rev. D*, 89(6):064048, March 2014. 1, 3.2.3.1

[136] P. B. Graff, A. Buonanno, and B. S. Sathyaprakash. Missing Link: Bayesian detection and measurement of intermediate-mass black-hole binaries. *Phys. Rev. D*, 92(2):022002, July 2015. 1, 3.2.3.1, 3.2.3.2, 5.2

[137] B. Miller, R. O'Shaughnessy, T. B. Littenberg, and B. Farr. Rapid gravitational wave parameter estimation with a single spin: Systematic uncertainties in parameter estimation with the SpinTaylorF2 approximation. *Phys. Rev. D*, 92(4):044056, August 2015. 1

[138] Michalis Agathos, Jeroen Meidam, Walter Del Pozzo, Tjonnie G. F. Li, Marco Tompitak, John Veitch, Salvatore Vitale, and Chris Van Den Broeck. Constraining the neutron star

equation of state with gravitational wave signals from coalescing binary neutron stars. *Phys. Rev.*, D92(2):023012, 2015. 1

[139] D. Trifirò, R. O'Shaughnessy, D. Gerosa, E. Berti, M. Kesden, T. Littenberg, and U. Sperhake. Distinguishing black-hole spin-orbit resonances by their gravitational wave signatures. II. Full parameter estimation. *Phys. Rev. D*, 93(4):044071, February 2016. 1

[140] Patricia Schmidt, Ian W. Harry, and Harald P. Pfeiffer. Numerical Relativity Injection Infrastructure. 2017. 3.2.2, 3.3.1

[141] C. R. Galley and P. Schmidt. Fast and efficient evaluation of gravitational waveforms via reduced-order spline interpolation. *ArXiv e-prints*, November 2016. 3.2.2

[142] V. Varma, P. Ajith, S. Husa, J. C. Bustillo, M. Hannam, and M. Pürrer. Gravitational-wave observations of binary black holes: Effect of nonquadrupole modes. *Phys. Rev. D*, 90(12):124004, 2014. 3.2.3.1, 3.4.1, 3.4.3.1, 3.4.3.2, 5.2

[143] J. Calderón Bustillo, S. Husa, A. M. Sintes, and M. Pürrer. Impact of gravitational radiation higher order modes on single aligned-spin gravitational wave searches for binary black holes. *Phys. Rev. D*, 93(8):084019, April 2016. 3.2.3.1, 3.2.3.1, 5.2

[144] Vijay Varma and Parameswaran Ajith. Effects of nonquadrupole modes in the detection and parameter estimation of black hole binaries with nonprecessing spins. *Phys. Rev.*, D96(12):124024, 2017. 3.2.3.1, 3.2.3.1, 3.2.3.1, 3.4.3.1, 3.4.3.2, 5.2

[145] J. Calderón Bustillo, P. Laguna, and D. Shoemaker. Detectability of gravitational waves from binary black holes: Impact of precession and higher modes. *Phys. Rev. D*, 95(10):104038, May 2017. 3.2.3.1, 3.2.3.1, 5.2

[146] D. A. Brown, A. Lundgren, and R. O'Shaughnessy. Nonspinning searches for spinning black hole-neutron star binaries in ground-based detector data: Amplitude and mismatch predictions in the constant precession cone approximation. *Phys. Rev. D*, 86(6):064020, September 2012. 3.2.3.1

[147] A. Lundgren and R. O'Shaughnessy. Single-spin precessing gravitational waveform in closed form. *Phys. Rev. D*, 89:044021, Feb 2014. 3.2.3.1

[148] Richard O'Shaughnessy, Prakash Nepal, and Andrew Lundgren. A semianalytic Fisher matrix for precessing BH-NS binaries. In *APS April Meeting Abstracts*, volume 2015 of *APS Meeting Abstracts*, page U13.006, April 2015. 3.2.3.1

[149] R. O'Shaughnessy, B. Vaishnav, J. Healy, and D. Shoemaker. Intrinsic selection biases of ground-based gravitational wave searches for high-mass black hole-black hole mergers. *Phys. Rev. D*, 82(10):104006, November 2010. 3.2.3.1

[150] C. Capano, Y. Pan, and A. Buonanno. Impact of higher harmonics in searching for gravitational waves from nonspinning binary black holes. *Phys. Rev. D*, 89(10):102003, May 2014. 3.2.3.1, 5.2

[151] C. Van Den Broeck and A. S. Sengupta. Phenomenology of amplitude-corrected post-Newtonian gravitational waveforms for compact binary inspiral: I. Signal-to-noise ratios. *Classical and Quantum Gravity*, 24:155–176, January 2007. 3.2.3.1

[152] R. N. Lang and S. A. Hughes. Measuring coalescing massive binary black holes with gravitational waves: The impact of spin-induced precession. *Phys. Rev. D*, 74(12):122001–+, December 2006. 3.2.3.1

[153] A. Klein, P. Jetzer, and M. Sereno. Parameter estimation for coalescing massive binary black holes with LISA using the full 2-post-Newtonian gravitational waveform and spin-orbit precession. *Phys. Rev. D*, 80(6):064027, September 2009. 3.2.3.1

[154] E. K. Porter and N. J. Cornish. Effect of higher harmonic corrections on the detection of massive black hole binaries with LISA. *Phys. Rev. D*, 78(6):064005, September 2008. 3.2.3.1

[155] R. N. Lang, S. A. Hughes, and N. J. Cornish. Measuring parameters of massive black hole binaries with partially aligned spins. *Phys. Rev. D*, 84(2):022002–+, July 2011. 3.2.3.1

[156] Tyson B. Littenberg, John G. Baker, Alessandra Buonanno, and Bernard J. Kelly. Systematic biases in parameter estimation of binary black-hole mergers. *Phys. Rev. D*, 87:104003, 2013. 3.4.1, 3.4.3.1, 3.4.3.2, 5.2

[157] LIGO Scientific Collaboration. LSC Algorithm Library. 2

[158] GWSurrogate. `https://pypi.python.org/pypi/gwsurrogate/`. 3.4.1

[159] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 2. The Waves' information and its extraction, with and without templates. *Phys. Rev.*, D57:4566–4587, 1998. 3.4.1

[160] Lee Lindblom, Benjamin J. Owen, and Duncan A. Brown. Model Waveform Accuracy Standards for Gravitational Wave Data Analysis. *Phys. Rev.*, D78:124020, 2008. 3.4.1

[161] Sean T. McWilliams, Bernard J. Kelly, and John G. Baker. Observing mergers of nonspinning black-hole binaries. *Phys. Rev.*, D82:024014, 2010. 3.4.1

[162] Katerina Chatziioannou, Antoine Klein, Neil Cornish, and Nicolas Yunes. Analytic Gravitational Waveforms for Generic Precessing Binary Inspirals. *Phys. Rev. Lett.*, 118(5):051101, 2017. 3.4.1

[163] Emily Baird, Stephen Fairhurst, Mark Hannam, and Patricia Murphy. Degeneracy between mass and spin in black-hole-binary waveforms. *Phys. Rev. D*, 87(2):024035, January 2013. 3.4.1

[164] Michele Vallisneri and Nicolas Yunes. Stealth Bias in Gravitational-Wave Parameter Estimation. *Phys. Rev.*, D87(10):102002, 2013. 3.4.1

[165] Marc Favata. Systematic parameter errors in inspiraling neutron star binaries. *Phys. Rev. Lett.*, 112:101101, 2014. 3.4.1

[166] B. P. Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119(14):141101, 2017. 3.4.2

[167] Chinmay Kalaghatgi, Mark Hannam, and Vivien Raymond. Parameter estimation with a spinning multimode waveform model. *Phys. Rev. D*, 101(10):103004, May 2020. 3.4.3, 3.4.4.4, 5.2

[168] Ethan Payne, Colm Talbot, and Eric Thrane. Higher order gravitational-wave modes with likelihood reweighting. *Phys. Rev. D*, 100(12):123017, December 2019. 3.4.3.3

[169] Manuela Campanelli, C. O. Lousto, and Y. Zlochower. Spinning-black-hole binaries: The orbital hang up. *Phys. Rev.*, D74:041501, 2006. 3.4.3.3

[170] K. Chatziioannou, G. Lovelace, M. Boyle, M. Giesler, D. A. Hemberger, R. Katebi, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi. Measuring the properties of nearly extremal black holes with gravitational waves. *Phys. Rev. D*, 98(4):044028, August 2018. 3.4.4.1

[171] B. Abbott et al. (The LIGO Scientific Collaboration and the Virgo Collaboration). Tests of general relativity with GW150914. *Phys. Rev. Lett*, 116:221101, May 2016. 3.4.4.2, 5.4

[172] Iara Ota and Cecilia Chirenti. Overtones or higher harmonics? Prospects for testing the no-hair theorem with gravitational wave detections. *Phys. Rev. D*, 101(10):104005, May 2020. 3.4.4.2

[173] Vijay Varma, Davide Gerosa, Leo C. Stein, François Hébert, and Hao Zhang. High-accuracy mass, spin, and recoil predictions of generic black-hole merger remnants. *Phys. Rev. Lett.*, 122(1):011101, 2019. 3.4.4.2

[174] Katerina Chatziioannou, Roberto Cotesta, Sudarshan Ghonge, Jacob Lange, Ken K. Y. Ng, Juan Calderón Bustillo, James Clark, Carl-Johan Haster, Sebastian Khan, Michael Pürrer, Vivien Raymond, Salvatore Vitale, Nousha Afshari, Stanislav Babak, Kevin Barkett, Jonathan Blackman, Alejandro Bohé, Michael Boyle, Alessandra Buonanno, Manuela Campanelli, Gregorio Carullo, Tony Chu, Eric Flynn, Heather Fong, Alyssa Garcia, Matthew Giesler, Maria Haney, Mark Hannam, Ian Harry, James Healy, Daniel

Hemberger, Ian Hinder, Karan Jani, Bhavesh Khamersa, Lawrence E. Kidder, Prayush Kumar, Pablo Laguna, Carlos O. Lousto, Geoffrey Lovelace, Tyson B. Littenberg, Lionel London, Margaret Millhouse, Laura K. Nuttall, Frank Ohme, Richard O'Shaughnessy, Serguei Ossokine, Francesco Pannarale, Patricia Schmidt, Harald P. Pfeiffer, Mark A. Scheel, Lijing Shao, Deirdre Shoemaker, Bela Szilagyi, Andrea Taracchini, Saul A. Teukolsky, and Yosef Zlochower. On the properties of the massive binary black hole merger GW170729. *Phys. Rev. D*, 100(10):104015, November 2019. 3.4.4.4, 5.4

[175] Michael Pürrer, Mark Hannam, and Frank Ohme. Can we measure individual black-hole spins from gravitational-wave observations? *Phys. Rev. D*, 93(8):084042, April 2016. 3.4.4.5

[176] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Davide Gerosa, Leo C. Stein, Lawrence E. Kidder, and Harald P. Pfeiffer. Surrogate models for precessing binary black hole simulations with unequal masses. *Physical Review Research*, 1(3):033015, October 2019. 3.5

[177] Juan Calderón Bustillo, Francesco Salemi, Tito Dal Canton, and Karan P. Jani. Sensitivity of gravitational wave searches to the full signal of intermediate-mass black hole binaries during the first observing run of advanced ligo. *Phys. Rev. D*, 97:024016, Jan 2018. 3.5

[178] B. P. Abbott, et. al., LIGO Scientific Collaboration, and Virgo Collaboration. Search for intermediate mass black hole binaries in the first and second observing runs of the Advanced LIGO and Virgo network. *Phys. Rev. D*, 100(6):064064, September 2019. 3.5.1

[179] Patricia Schmidt, Frank Ohme, and Mark Hannam. Towards models of gravitational waveforms from generic binaries: Ii. modelling precession effects with a single effective precession parameter. *Phys. Rev. D*, 91:024043, Jan 2015. 3.5.1

[180] A. Sesana. Prospects for Multiband Gravitational-Wave Astronomy after GW150914. *Physical Review Letters*, 116(23):231102, June 2016. 3.5.1

[181] Karan Jani, Deirdre Shoemaker, and Curt Cutler. Detectability of intermediate-mass black holes in multiband gravitational wave astronomy. *Nature Astronomy*, 4:260–265, November 2019. 3.5.1

[182] Manuel Arca Sedda, Christopher Berry, Karan Jani, Pau Amaro-Seoane, Pierre Auclair, Jonathon Baird, Tessa Baker, Emanuele Berti, Katelyn Breivik, Chiara Caprini, Xian Chen, Daniela Doneva, Jose M. Ezquiaga, K. E. Saavik Ford, Michael Katz, Shimon Kolkowitz, Barry McKernan, Guido Mueller, Germano Nardini, Igor Pikovski, Surjeet Rajendran, Alberto Sesana, Lijing Shao, Nicola Tamanini, Niels Warburton, Helvi Witek, Kaze Wong, and Michael Zevin. The missing link in gravitational-wave astronomy: Discoveries waiting in the decihertz range, 2019. 3.5.1

[183] S. Vitale. Multiband Gravitational-Wave Astronomy: Parameter Estimation and Tests of General Relativity with Space- and Ground-Based Detectors. *Physical Review Letters*, 117(5):051102, July 2016. 3.5.1

[184] Curt Cutler, Emanuele Berti, Kelly Holley-Bockelmann, Karan Jani, Ely D. Kovetz, Shane L. Larson, Tyson Littenberg, Sean T. McWilliams, Guido Mueller, Lisa Randall, Jeremy D. Schnittman, David H. Shoemaker, Michele Vallisneri, Salvatore Vitale, and Kaze W. K. Wong. What can we learn from multi-band observations of black hole binaries? *Bulletin of the AAS*, 51(3):109, May 2019. 3.5.1

[185] LIGO Scientific Collaboration and Virgo Collaboration. GW190521.1: A Binary Black Hole Coalescence with a Total Mass of 150 $M_{\odot}$. Technical Report P2000020, LSC, 2020. 4.1, 4.2, 4.2.1, 4.2.6, 5.4

[186] Tyson B. Littenberg and Neil J. Cornish. Bayesian inference for spectral estimation of gravitational wave detector noise. *Phys. Rev. D*, 91(8):084034, April 2015. 4.2

[187] Katerina Chatziioannou, Carl-Johan Haster, Tyson B. Littenberg, Will M. Farr, Sudarshan Ghonge, Margaret Millhouse, James A. Clark, and Neil Cornish. Noise spectral

estimation methods and their impact on gravitational wave measurement of compact binary mergers. *Phys. Rev. D*, 100(10):104004, November 2019. 4.2

[188] The LIGO Scientific Collaboration and The Virgo Collaboration. Binary black hole population properties inferred from the first and second observing runs of advanced LIGO and advanced virgo. *The Astrophysical Journal*, 882(2):L24, sep 2019. 4.2.1

[189] B. P. Abbott, et. al., (The LIGO Scientific Collaboration, and The Virgo Collaboration). GW190425: Observation of a Compact Binary Coalescence with Total Mass ~ 3.4 $M_{\odot}$. *Astrophysical Journal*, 892(1):L3, March 2020. 4.2.1

[190] Duncan A. Brown, Prayush Kumar, and Alexander H. Nitz. Template banks to search for low-mass binary black holes in advanced gravitational-wave detectors. *Phys. Rev. D*, 87(8):082004, April 2013. 5.2

[191] Ian Harry, Juan Calderón Bustillo, and Alex Nitz. Searching for the full symphony of black hole binary mergers. *Phys. Rev.*, D97(2):023004, 2018. 5.2

[192] Prayush Kumar, Jonathan Blackman, Scott E. Field, Mark Scheel, Chad R. Galley, Michael Boyle, Lawrence E. Kidder, Harald P. Pfeiffer, Bela Szilagyi, and Saul A. Teukolsky. Constraining the parameters of GW150914 and GW170104 with numerical relativity surrogates. *Phys. Rev.*, D99(12):124005, 2019. 5.2

[193] B. P. Abbott et al. GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence. *Astrophys. J.*, 851(2):L35, 2017. 5.4

[194] Champion, Benjamin and Birnholtz, Ofek and O'Shuaghnessy, Richard. Multi-Messenger Astrophysics Parameter Estimation for GW and EM data channels. Technical Report G1900532, LSC, 2019. 5.4

[195] M. Breschi, R. O'Shaughnessy, O. Birnholz, and J. Lange. Imr consistency tests with higher modes on o2 events. *MS Thesis; Available as LIGO-L1800348 and arxiv:1903.05982*, 2018. 5.4

[196] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al. Tests of General Relativity with GW170817. *Phys. Rev. Lett*, 123(1):011102, July 2019. 5.4

[197] B. P. Abbott, et. al., (The LIGO Scientific Collaboration, and The Virgo Collaboration). Tests of general relativity with the binary black hole signals from the LIGO-Virgo catalog GWTC-1. *Phys. Rev. D*, 100(10):104036, November 2019. 5.4

[198] Jam Sadiq, Yosef Zlochower, Richard O'Shaughnessy, and Jacob Lange. Hybrid waveforms for generic precessing binaries for gravitational-wave data analysis. *Phys. Rev. D*, 102(2):024012, July 2020. 5.5

[199] Pau Amaro-Seoane, Heather Audley, Stanislav Babak, John Baker, Enrico Barausse, Peter Bender, Emanuele Berti, Pierre Binetruy, Michael Born, Daniele Bortoluzzi, Jordan Camp, Chiara Caprini, Vitor Cardoso, Monica Colpi, John Conklin, Neil Cornish, Curt Cutler, Karsten Danzmann, Rita Dolesi, Luigi Ferraioli, Valerio Ferroni, Ewan Fitzsimons, Jonathan Gair, Lluis Gesa Bote, Domenico Giardini, Ferran Gibert, Catia Grimani, Hubert Halloin, Gerhard Heinzel, Thomas Hertog, Martin Hewitson, Kelly Holley-Bockelmann, Daniel Hollington, Mauro Hueller, Henri Inchauspe, Philippe Jetzer, Nikos Karnesis, Christian Killow, Antoine Klein, Bill Klipstein, Natalia Korsakova, Shane L Larson, Jeffrey Livas, Ivan Lloro, Nary Man, Davor Mance, Joseph Martino, Ignacio Mateos, Kirk McKenzie, Sean T McWilliams, Cole Miller, Guido Mueller, Germano Nardini, Gijs Nelemans, Miquel Nofrarias, Antoine Petiteau, Paolo Pivato, Eric Plagnol, Ed Porter, Jens Reiche, David Robertson, Norna Robertson, Elena Rossi, Giuliana Russano, Bernard Schutz, Alberto Sesana, David Shoemaker, Jacob Slutsky, Carlos F. Sopuerta, Tim Sumner, Nicola Tamanini, Ira Thorpe, Michael Troebs, Michele Vallisneri, Alberto Vecchio, Daniele Vetrugno, Stefano Vitale, Marta Volonteri, Gudrun Wanner, Harry Ward, Peter Wass, William Weber, John Ziemer, and Peter Zweifel.

Laser Interferometer Space Antenna. *arXiv e-prints*, page arXiv:1702.00786, Feb 2017. 5.5